Proceedings of:

The Sixth Annual Conference

The American Society for Precision Engineering

October 13-18, 1991

Sweeney Convention Center

Santa Fe, New Mexico

The American Society for Precision Engineering (ASPE) is a multidisciplinary professional and technical society concerned with research and development, design, manufacture and measurement of high accuracy components and systems. ASPE activities encompass relevant aspects of mechanical, electronic, optical and production engineering, physics, chemistry, and computer and materials science. Membership is open to anyone interested in any aspect of precision engineering.

Founded in 1986, ASPE provides a new focus for a diverse but important community. Other professional organizations have covered aspects of precision engineering, always as a sideline to their principal goals. ASPE is based on the core of generic concepts necessary to achieve precision in any application; independent of discipline, ASPE intends to be the focus for precision technology —and to represent all facets from research to application.

ASPE — The American Society for Precision Engineering
401 Oberlin Road, Suite 108
P.O. Box 10826
Raleigh, NC 27605-0826
Phone (919) 839-8444 Fax (919) 839-8039

These are exciting times for precision engineering. Scientific advances in materials, metrology, machine design, and controls have extended the potential for design and manufacturing into the regime of nanotechnology. The 90's will bring an expanded, integrated view of precision manufacturing with increased emphasis on metrology and material forming processes. Driven by applications for mechanical parts (bearings and gears), electronic components (VLSI and electro-optical devices), and optical assemblies, there will be more systems which will require improved understanding of metrology, machining, material response, sensors, control, and machine design to achieve their goals. The aim of the American Society for Precision Engineering is to serve as a focus to emphasize the need for information exchange in this important area.

We are proud to present the program for the Sixth Annual Conference of the ASPE. This international forum for the exchange of ideas, for the discussion of solutions to problems, and for the presentation of results involving precision engineering involves professionals in universities, government laboratories, and private industry. The meeting offers the opportunity for multidisciplinary interaction and a view of the breadth of precision engineering.

This program reflects the vitality of the ASPE and the zest of its Annual Conference.

Welcome to Santa Fe.

1991 ASPE Conference Organizing Committee

Richard L. Rhorer, Chair
Los Alamos National Laboratory

Torsten Bispink, Fraunhofer-IPT
Norman Bobroff, IBM
Len E. Chaloux, Rank Taylor Hobson, Inc.
Thomas A. Dow, NC State University
W. Tyler Estler, Nat'l Institute of Standards & Technology
Robert J. Hocken, UNC-Charlotte
Fred Jones, Martin Marietta Energy Systems
Don A. Lucca, Oklahoma State University
George R. Platt, Texas Instruments
Ronald O. Scattergood, NC State University
Alexander Slocum, Massachusetts Institute of Technology
Margaret Stedman, National Physical Laboratory
E. Clayton Teague, Nat'l Institute of Standards & Technology
Daniel C. Thompson, Lawrence Livermore National Laboratory
Carl A. Zanoni, Zygo Corporation

1991 Annual Conference

Contents

Foreword — Organizing Committee ..... *ii*

Preface — Corporate Members ..... *iv*

Cooperating Sponsors — Technical Exhibitors ..... *v*

Program ..... *vi*

Commercial/Forum/Distinguished Lecturer *1*

Technical Paper Index *2-6*

Technical Papers *7-225*

Author Index *226-227*

Preface

This book comprises the proceedings of the 1991 Annual Conference. The contributions reflect the authors' opinions and are published as presented to ASPE, without change. Their inclusion in this publication does not necessarily constitute endorsement by the ASPE, or its editorial staff.

Individual readers of this book and nonprofit libraries acting for them are permitted to make fair use of the material contained in this book to be used in teaching or research; provided they copy said paper in total, as written by the author. Permission is granted to quote excerpts from the scientific or technical works in this book with proper acknowledgment of the source to include: the author's name, the book name, and the ASPE tag line. Reproduction of figures and tables is likewise permitted from the contents of this book, provided that the same acknowledgment of the same source and content is printed.

Any other type reproduction, systematic copying or multiple reproduction of any material in this book is expressly prohibited except with the written permission of the ASPE editorial staff, addressed to the Society's headquarters.

Printed in the United States of America
AMERICAN SOCIETY FOR PRECISION ENGINEERING

ISBN 1-887706-06-2

Corporate Members of the ASPE

NAME	CONTACT	LOCATION
3M Company	William J. Bryan	Petaluma, CA
Chardon Tool & Supply Co., Inc.	Weldon Bennett	Chardon, OH
Cranfield Precision Engineering Ltd.	Patrick McKeown	Cranfield, ENGLAND
Dover Instrument Corporation	John R. Hero	Westboro, MA
GCA Tropel	Donald P. McClimans	Fairport, CT
IBM Corporation	Norman Bobroff	Yorktown Hts., NY
Micro Encoder, Inc.	Kim Atherton	Kirkland, WA
Moore Special Tool	Michael L. Majlak	Bridgeport, CT
Nissin Machine Works, Ltd.	Mitsuaki Takahashi	Hamamatsu, JAPAN
Norton Company	Francis D. Renzoni	Arden, NC
Professional Instruments Co.	Ted J. Arneson, Jr.	Minneapolis, MN
Rank Taylor Hobson/Rank Pneumo	John B. Byrd, III	Keene, NH
Tencor Instruments	Graham Siddall	Mountain View, CA
Zygo Corporation	Carl Zanoni	Middlefield, CT

Cooperating Sponsors

Brown & Sharpe

Laser Focus World

Photonics Spectra

Rank Taylor Hobson/Rank Pneumo

Technical Exhibitors

ADE Corporation
A. G. Davis Gage & Engineering Co.
Alson E. Hatheway, Inc.
Brown & Sharpe
Chapman Instruments, Inc.
Chardon Tool & Supply Co., Inc.
Cranfield Precision Engineering
Hewlett-Packard Company
Lion Precision
Micromap Corporation
NCSU Precision Engineering Center
Norton Company
Precision Systems
Professional Instruments Co.
Rank Taylor Hobson/Rank Pneumo
Speedring Systems, Inc.
Surface Finishes Co., Inc.
U.S. Contour Tooling, Inc.
University of North Carolina at Charlotte
Vermont Photonics
Zygo Corporation

Program

1991 ASPE Annual Conference

Time	Sun., Oct. 13	Mon., Oct. 14	Tues., Oct. 15	Wed., Oct. 16	Thurs., Oct. 17	Fri., Oct. 18
8:00						
9:00	Tutorial	Tutorial	Plenary Session	Technical Session 3	Technical Session 5	Industrial Visits: LANL 8:00-4:00 Albuquerque 8:00-3:00
10:00			Break	Break	Break	
11:00			Technical Session 1	Technical Session 4	Technical Session 6	
12:00			Lunch + Committee Meetings	Lunch + Poster Session	Lunch + Poster Session	
1:00						
2:00	Tutorial	Tutorial	Technical Session 2	Business Meeting	Technical Session 7	
3:00			Break	OPEN	Break	
4:00			Commercial Session		Technical Session 8	
5:00						
6:00	Board Meeting					
7:00		Welcome Reception	Distinguished Lecture	Open Forum	Museum Tour and Dinner	
8:00			Hospitality Suite			
9:00				Hospitality Suite		
10:00						

Other Technical Sessions

Commercial Session

Tuesday, October 15, 3:30-5:00

Session Chair: Leonard E. Chaloux, Rank Taylor Hobson/Rank Pneumo

Company representatives are invited to make brief presentations on products associated with precision engineering during this special afternoon session. The intent of the Commercial Session is to introduce the ASPE community to limitations on performance of existing approaches, new approaches relevant to precision engineering and new advances that can be expected in the future.

Open Forum Session

Wednesday, October 16, 7:00 - 9:00

Session Chair: Richard Rhorer, Los Alamos National Laboratory

This evening is devoted to short presentations of current problems, and new and significant results and conclusions in precision engineering. Sign-up sheets for interested presenters will be available at the Conference Registration Table.

Distinguished Lecturer

Tuesday, October 15, 7:00

James B. Bryan

Jim Bryan has devoted his entire career to precision engineering. Making and measuring precision components and identifying the enemies of ideal performance were his goal. He became an apostle of the Abbe comparator principle and high priest for controlling temperature. An active member of B-89 since its 1963 founding, he was the father of three of its major standards. His key role in the development of precision diamond turning is world famous.

Always generous with help to others, he has organized innumerable seminars, prepared dozens of papers, and served on the advisory panel of the Center for Manufacturing Technology of NBS. He has been honored many times, including medals from SME and SPIE. Jim is an active member of CIRP, ASME, and SME, and served on the board of directors of the American Society for Precision Engineering.

Since retiring from Lawrence Livermore National Laboratory, he has become a busy consultant when not sailing in the San Francisco Bay.

Technical Paper Index

The technical program for the 1991 ASPE Annual Conference contains reports on current research from China, Europe, Japan and the US. Authors from national laboratories, universities, and industry will share their conclusions on a wide variety of topics ranging from gravitational wave detection to design of drive systems for high accuracy rotary positioning devices.

In addition to nine technical sessions, there will be an open forum where recent results can be discussed, a commercial session where equipment manufacturers will give an update on their recent innovations, and two poster sessions.

Plenary Session
Precision Engineering

Tuesday, October 15, 8:30 - 10:00 — *Session Chair: Robert Parks (ASPE President)*

Invited Papers:

1. ***The Metrology Associated with Interferometric Gravitational Wave Detection***
   R. Weiss (Massachusetts Institute of Technology) ..... NA
2. ***State-Of-The-Art and Ultimate Physical Limitations of Dimensional Metrology***
   H. Kunzmann (Physikalisch-Technische Bundesanstalt) ..... page 9
3. ***Servo-Guided Planar Stages and Other X-Y Stages for Microlithography: A Review***
   R. A. Kendall (IBM Corp.) ..... page 13

Session I
Dimensional Metrology

Tuesday, October 15, 10:30 - 12:00 — *Session Chair: Norman Bobroff (IBM T. J. Watson Research Center)*

1. ***Application and Performance of Capacitance Micrometry as a Super-Precision Transfer Standard***
   S. Harb, D.G. Chetwynd and S.T. Smith (University of Warwick) ..... page 17
2. ***Advanced Angle Metrology at the National Institute of Standards and Technology***
   W.T. Estler, Y.H. Queen and D. Gilsinn (National Institute of Standards and Technology) and D. Pieczulewski (A.G. Davis Gage & Engineering Co.) ..... page 21
3. ***Scales vs. Interferometers: A Comparison of Metrology Systems***
   G. Wyntjes, D. Mitchell and M. Hercher (OPTRA Inc.) ..... page 25
4. ***Scanning Probe Microscopes as Co-ordinate Measuring Machines***
   M. Stedman (National Physical Laboratory) ..... page 29

Session II
Machine Metrology

Tuesday, October 15, 1:30 - 3:00 — *Session Chair: Margaret Stedman (National Physical Laboratory)*

1. ***The Characterization of Analog Probes for Use on Coordinate Measuring Machines***
   H. Harary (National Institute of Standards and Technology) and D. Clermont, C. Lavigne, P. Chollet and A. Clement (Institut Superieur des Materiaux et de la Construction Mecanique) ..... *page 32*
2. ***Static and Dynamic Properties of Free Standing Ball Bars***
   S.D. Phillips, B. Borchardt and T.D. Doiron (National Institute of Standards and Technology) and J. Henry (Sheffield Measurement Corporation) ..... *page 33*
3. ***The Efficiency of Artifact Based Procedures to Calibrate Coordinate Measuring Machines***
   J.A. Soons and P.H. Schellekens (Eindhoven University of Technology) ..... *page 37*
4. ***Error Analysis for X-Y Interferometric Metrology***
   N. Bobroff (IBM T.J. Watson Research Center) ..... *page 43*

Session III
Interferometry

Wednesday, October 16, 8:30 - 10:00 — *Session Chair: Carl A. Zanoni (Zygo Corporation)*

1. ***Electronic Limitations in Phase Meters for Heterodyne Interferometry***
   N. Oldham, P. Hetrick, J. Kramar, W. Penzes, T. Wheatley and C. Teague (National Institute of Standards and Technology) ..... *page 47*
2. ***An Interferometric Instrument for the Metrology of Micromechanisms***
   L.P. Howard and R.J. Hocken (University of North Carolina at Charlotte) ..... *page 50*
3. ***Some Observations on the Performance of Commercial Phase Measuring (Fizeau) Interferometers Used in Surface Figure Metrology***
   C.J. Evans and W.T. Estler (National Institute of Standards and Technology) ..... *page 54*
4. ***A Focused View on Aspherics***
   D.M.G. Stevens (Rank Taylor Hobson Ltd.) ..... *page 58*

Session IV
Controls and Precision Manufacturing

Wednesday, October 16, 10:30 - 12:00 — *Session Chair: Dan E. Luttrell (Rank Taylor Hobson, Inc.)*

1. ***Controller Design for a Modern Diamond Turning Machine***
   M.H. Miller, K.P. Garrard, T.A. Dow and L.W. Taylor (North Carolina State University) ..... *page 62*
2. ***Nanometer Motion Control via Magnetic Suspension***
   D.L. Trumper (University of North Carolina at Charlotte) and A.H. Slocum (Massachusetts Institute of Technology) ..... *page 66*
3. ***Diamond Point Machined Snap-Together Optics***
   J.P. Schaefer (Texas Instruments) ..... *page 70*
4. ***The Effects of Fundamental Axial Spindle Motion on the Optical Performance of Diamond Turned Surfaces***
   M.C. Gerchman (Rank Taylor Hobson, Inc.) ..... *page 74*

Session V
Precision Drives

Thursday, October 17, 8:30 - 10:00 — *Session Chair: Daniel C. Thompson (Lawrence Livermore National Laboratory)*

1. ***Development of Drive-Systems for High Accuracy Rotary Positioning Devices***
   M. Weck and J. Luderich (Fraunhofer-Institut für Produktionstechnologie) ..... *page 76*
2. ***Feed and Frictional Load Characteristics of Carriage on Plain Bearing Guideways in the Range of Nanometer Positioning***
   A. Kanai and M. Miyashita (Ashikaga Institute of Technology) and J. Yoshioka (NUTR K.K.) ..... *page 80*
3. ***Ultra-precise AC Linear Motor for Nanometer Position Control***
   S. Futami and A. Furutani (Yaskawa Electric Corporation) ..... *page 85*
4. ***Model Reference Adaptive Control of Nonlinear Microdynamics of a Ball-Screw Driven Precision Slide***
   P.I. Hubbel and P.I. Ro (North Carolina State University) ..... *page 89*

Session VI
Ultraprecision Fabrication of Ductile Materials

Thursday, October 17, 10:30 - 12:00 — *Session Chair: Don A. Lucca (Oklahoma State University)*

1. ***Aluminum Optics, But of Which Alloy and What Temper?***
   R.S. Polvani and C.J. Evans (National Institute of Standards and Technology) ..... *page 93*
2. ***Single Point Diamond Machining Induced Damage on Single Crystal Copper, Aluminum and Gold***
   J.W. Carr and C. Narayan (IBM T.J. Watson Research Center) and J. Kim (Goldstar Electronics) ..... *page 96*
3. ***Molecular Dynamics Simulation of the Chip Forming Process in Single Crystal Copper and Comparison with Experimental Data***
   I.F. Stowers and J.F. Belak (Lawrence Livermore National Laboratory); D.A. Lucca and R. Komanduri (Oklahoma State Univ.); R.L. Rhorer (Los Alamos National Laboratory); T. Moriwaki (Kobe University); K. Okuda (Kobe Technical College); N. Ikawa, S. Shimada and H. Tanaka (Osaka University); T.A. Dow and J.D. Drescher (North Carolina State University) ..... *page 100*
4. ***Machining Forces in Diamond Turning of Copper***
   J.D. Drescher and T.A. Dow (North Carolina State University) ..... *page 104*

Session VII
Ultraprecision Fabrication of Brittle Materials

Thursday, October 17, 2:00 - 3:30 — *Session Chair: Jeffrey W. Carr (IBM T.J. Watson Research Center)*

1. ***Analysis and Modeling of Diamond Turning of Brittle Materials***
   M. Tidwell and R.O. Scattergood (North Carolina State University) ..... *page 108*
2. ***Material Removal Mechanisms Involved in the Single-Point Diamond Turning of Brittle Materials***
   C-L. Chao (Chung-Cheng Institute of Technology) and A.E. Gee (Cranfield Institute of Technology) ..... *page 112*
3. ***Fixed-Abrasive Grinding of CVD SiC to Optical Quality***
   T. Bifano and Y. Yi (Boston University) ..... *page 116*
4. ***Wheel Speed Effects in Precision Contour Grinding***
   S.C. Fawcett and T.A. Dow (North Carolina State University) ..... *page 120*

Session VIII
Ultraprecision Fabrication of Glass

Thursday, October 17, 4:00 - 5:30 — *Session Chair: Ronald O. Scattergood (North Carolina State University)*

1. ***Experimental Machining of Glass to Optical Finishes in a Closed Loop Configuration***
   C.G. Blough, E.G. Loewen and D.T. Moore (University of Rochester) ..... page 124
2. ***Grinding of Optical Glass with Loose Polycrystalline Synthetic Diamond Abrasives***
   M.J. Cumbo (Bausch and Lomb) and A. Lindquist, S.B. Ng, T. Rich, Y. Sabharwal and S.D. Jacobs (University of Rochester) ..... page 128
3. ***Shear Mode Grinding of BK7 Glass and Evaluating Ground Surface***
   J. Takeda, K. Imai and H. Hashimoto (Kanagawa Institute of Technology) ..... page 132
4. ***A Mirror-Finish Slicing Technology with Electrolytic Dressing for Hard and Brittle Materials***
   T. Kanda and M. Mitsuhashi (NEC Corp.) ..... page 133

Poster Session I
Metrology and Interferometry

Wednesday, October 16, 12:00 - 2:00 — *Session Chair: Robert D. Day (Los Alamos National Laboratory)*

1. ***A Long Range STM/AFM***
   C. Lavigne, J.A. Miller and R.J. Hocken (University of North Carolina at Charlotte) ..... page 137
2. ***Surface Characterisation via Optical Diffraction***
   D. Mansfield (Rank Taylor Hobson Ltd.) ..... page 141
3. ***The Measurement Variation of Surface Profile Parameters: A Viewpoint from Standardisation***
   T.Y. Lin, P.J. Sullivan and K.J. Stout (University of Birmingham) and K.L. Ko (Industrial Technology Research Institute) ..... page 145
4. ***Simultaneous Measurement of Spindle Error Motions: A Simple and Economical Method***
   S. Kumaran, T. Fujii and R.J. Hocken (University of North Carolina at Charlotte) ..... page 149
5. ***Using Scanning Probe Microscopes for Examining Precision Machined Tools and Machined Surfaces***
   M.R. Rodgers and V.B. Elings (Digital Instruments, Inc.) ..... page 153
6. ***Comparison of Angle Measuring Interferometers and Autocollimators***
   P. Greier and N. Bobroff (IBM T.J. Watson Research Center) and W.T. Estler and Y.H. Queen (National Institute of Standards and Technology) ..... page 157
7. ***The Characterization of Multiprocess Surfaces (Is This a Case for a Paradigm Shift?)***
   P.J. Scott (Rank Taylor Hobson Ltd.) ..... page 161
8. ***On-Machine Measurement of Surface Profile by Fiber Optic Heterodyne Interferometry***
   K. Sakaya and T. Ohura (Industrial Research Institute of Ishikawa) ..... page 165
9. ***A Hierarchical 2D Phase Unwrapping Strategy Immune to Spike Noise and Discontinuities of Larger Scale***
   T.R. Judge and P.J. Bryanston-Cross (University of Warwick) ..... page 169
10. ***Evaluation of Topography of Precision Engineering Part Surfaces***
    J. Mandak (Slovak Metrological Society) ..... page 173

Poster Session II
Machining Design & Controls and Ultraprecision Fabrication

Thursday, October 17, 12:00 - 2:00 — *Session Chair: Philip R. Hannah (Consultant)*

1. ***Development of a New Vibration Reduction Control Scheme for Precision Slide Motion***
   J.A. Abler and P.I. Ro (North Carolina State University) ..... *page 177*
2. ***The Preloaded Ball Screw as a Nanometer Actuator***
   J.F. Cuttino and T.A. Dow (North Carolina State University) ..... *page 181*
3. ***Design of Three Groove Kinematic Couplings***
   A.H. Slocum (Massachusetts Institute of Technology) ..... *page 185*
4. ***A Fundamental Structure for Intelligent Manufacturing***
   Y. Hatamura, T. Nagao and M. Mitsuishi (University of Tokyo) ..... *page 188*
5. ***Statistical Concept of System Functional Tolerancing Including Operational Conditions***
   V.M. Izraelev (3M Storage Media Laboratory) ..... *page 192*
6. ***Microtilt Controlled Chuck Faceplate for Ultra-precise Single-point Diamond Machining***
   A.E. Gee, J. Duduch, H. Mizuno, Y. Hara, H. Aoyama and K. Flynn (Cranfield Institute of Technology), C-L. Chao (Chung-Cheng Institute of Technology), K.E. Puttick and M. Rudman (University of Surrey) and S.G. McCandlish (University of Exeter) ..... *page 194*
7. ***Improvement of the Precision of Machining by the Use of Inertial and Gravitational Drives***
   E.S. Geskin and M.A. Shimanovich (New Jersey Institute of Technology) ..... *page 198*
8. ***Precision Grinding of Thin Quartz Wafers***
   J. Hosler and T. Bifano (Boston University) ..... *page 200*
9. ***Short-Crack Toughness Determination for Brittle Materials***
   S.M. Smith and R.O. Scattergood (North Carolina State University) ..... *page 205*
10. ***Microfabrication of Ultraprecision Structures Using Focused Ion Beam Machining***
    S.T. Davies (University of Warwick) ..... *page 209*
11. ***Environmental Effects on Subsurface Damage in Glass Machining***
    D. DePiero and T. Bifano (Boston University) ..... *page 213*
12. ***Influence of the Quality of Natural Diamond on Tool Wear in Precision Turning***
    M. Sawa, M. Masuda, Y. Maeda and R. Itoo (Hitachi, Ltd.) ..... *page 217*
13. ***Evaluation of Environmentally Safe Cleaning Agents on Diamond Turned Optics***
    L.A. Theye and R.D. Day (Los Alamos National Laboratory) and R. Schubert (Bell Communications Research) ..... *page 222*

Technical Papers

STATE-OF-THE-ART AND ULTIMATE PHYSICAL LIMITATIONS OF DIMENSIONAL METROLOGY

Horst Kunzmann

Physikalisch-Technische Bundesanstalt
Braunschweig, Germany

The objective of dimensional metrology is to provide quantified information about surface texture, form and shape, dimension, distances between positions of machine parts, angles between orientation etc..

When highest accuracy is the target to be met, any solution of a measuring problem is the accomplishment of either one task, or a combination of different tasks.
These major tasks of dimensional metrology are:

- realization of the unit of length or angle
- measurement of translation or rotation
- transfer of the measurands from the reference system to the test piece

At the same time, it is absolutely necessary:

- to measure the temperature
- to measure the refractive index of the air

and

- to know the material properties of the instrument and of the test piece
- to optimize the measurement conditions

Concerning dimensional metrology three physical units of the SI-System are involved. That are:

- the meter for length
- the radian for the plane angle
- the Kelvin resp. °Celsius for the temperature

The best realization of the unit of length is given by the vacuum wavelength of the Iodine stabilized He-Ne-Laser ($\lambda_0 \approx 0.633$ µm) with an relative accuracy of $u\lambda_0 = 10^{-10}\,\lambda_0$ (see Fig. 1).

The realization of the radian as the unit for the plane angle is without physical and technical relevance and has never been undertaken; the angle of a full circle is correct by definition!

The temperature influences directly dimensional metrology as all materials have a thermal coefficient of expansion, either smaller or greater. Measurement results are only correct given for the reference temperature t_{ref} = 20°C by international agreement in 1931. The realization of the reference temperature and temperature measurements near 20°C can be done with an ultimate accuracy of appr. 0.1 mK. But it should be noted a change in the International Practical Temperature Scale on January 1990, which increased the reference temperature by a value of 5 mK (see Fig. 2). Dimensional changes in the order of sine 10^{-8} were the consequence.

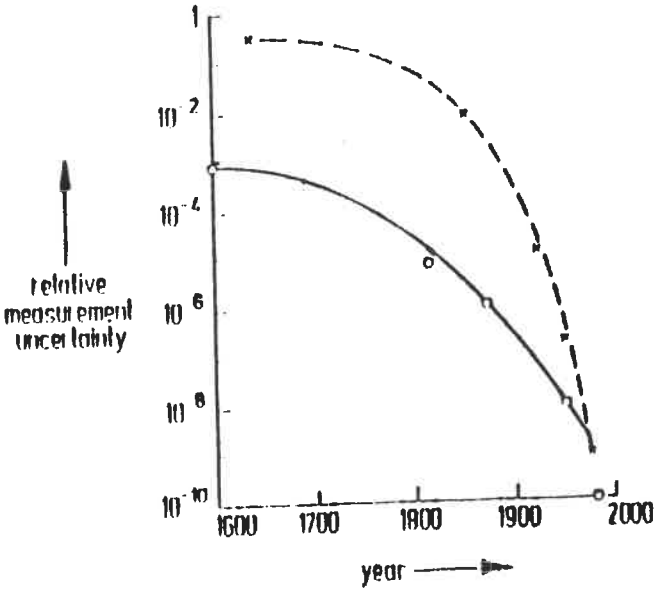


Fig.1
Improvement of accuracy for the realization of unit of length and the speed of light in vacuum (after Petley, NPL, 1987)

——— unit of length

------ speed of light

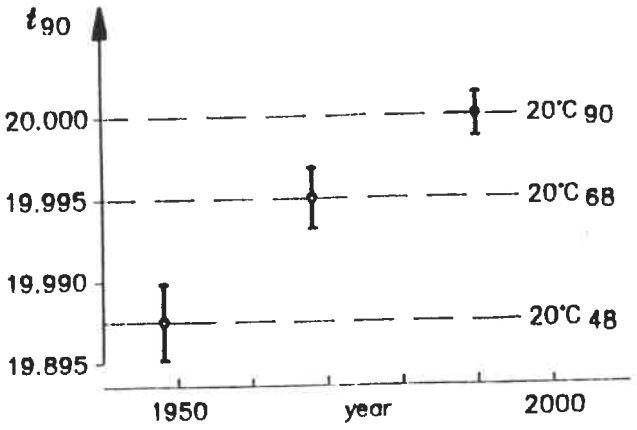


Fig.2
Changes of the reference temperature 20 °C

Besides these ultimate limitations given by the uncertainty of the realization of the involved physical units,

the physical principles, and

the perfection of the metrological control of the technologies,

applied to solve a measurement problem, set the limits to be taken into account.

Angle measurements have been since centuries a field of interest for astronomy, geodesy and precision engineering. As there is no unit realization setting a limitation, the technology itself set the ultimate state of the art. Principles of self-calibration, error reduction by randomization of systematic errors, and an avaraging of residual errors were applied, based on the technology available and with increasing accuracy (see Fig. 3).

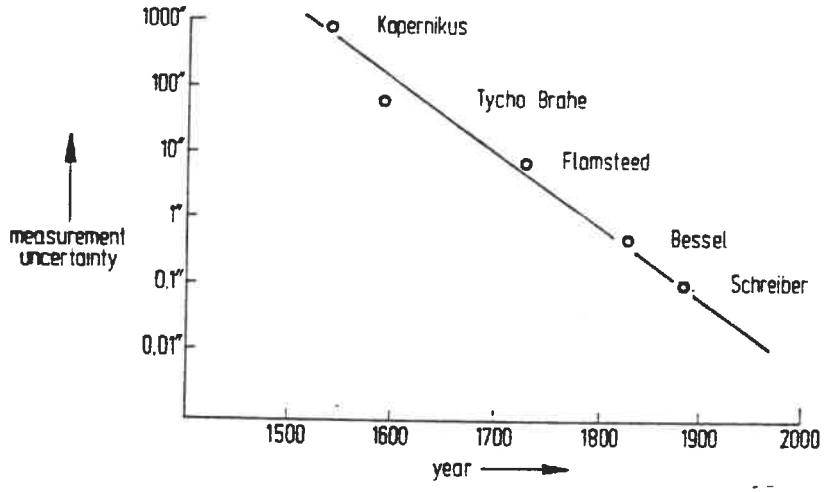


Fig.3 Improvement of accuracy for angle measurements (after C. Evans, NIST, 1987)

Linear displacement measurements of highest accuracy is done by laserinterferometry, today. Fig.4 outlines the accuracy limits in absolute and relative dimensions as a function of displacement length.

The most relevant limitations are given by the perfection of fractioning the interference orders and the grade of optimization of the measurement conditions concerning the refractive index of the air.

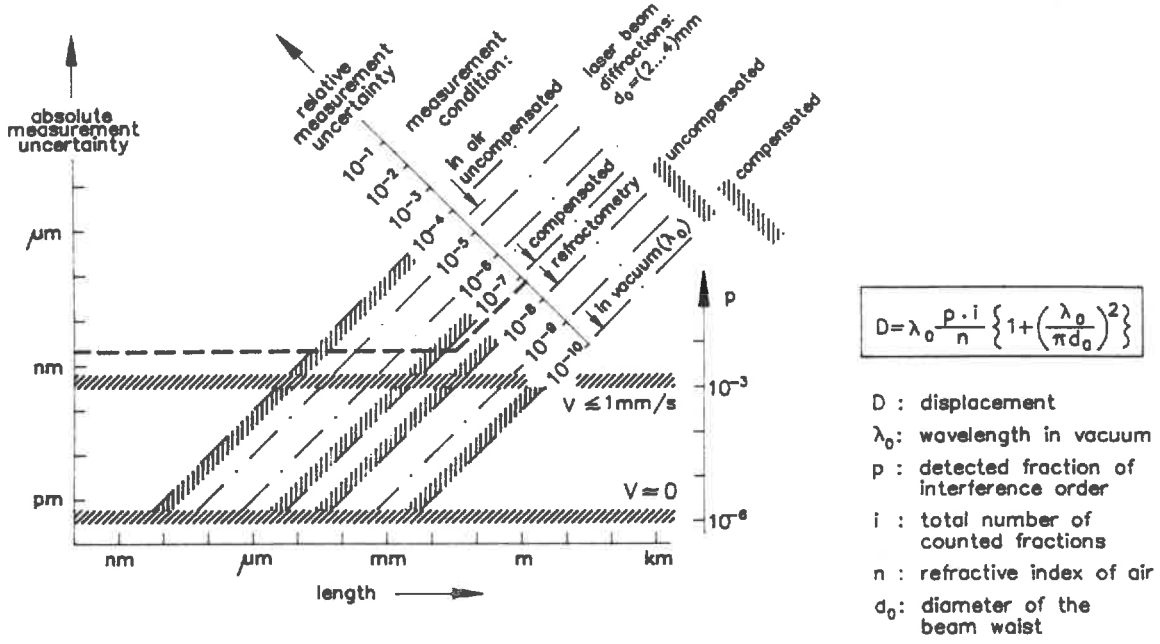


Fig.4 Accuracy limits for laserinterferometric displacement measurements

For microtopography, form and shape measurements, and the detection of microstructures there are probing systems available with a sensitivity and repeatability down to the 0.01 nm range. However, as the physical interaction between the surface and the different probing processes are not known with the same accuracy, about 10...20 nm must be assumed to be the limit of accuracy of dimensional measurements for those material standards, for which the dimensions are given by surfaces of opposite direction. Fig. 5 summarizes the situation for dimensional metrolgy.

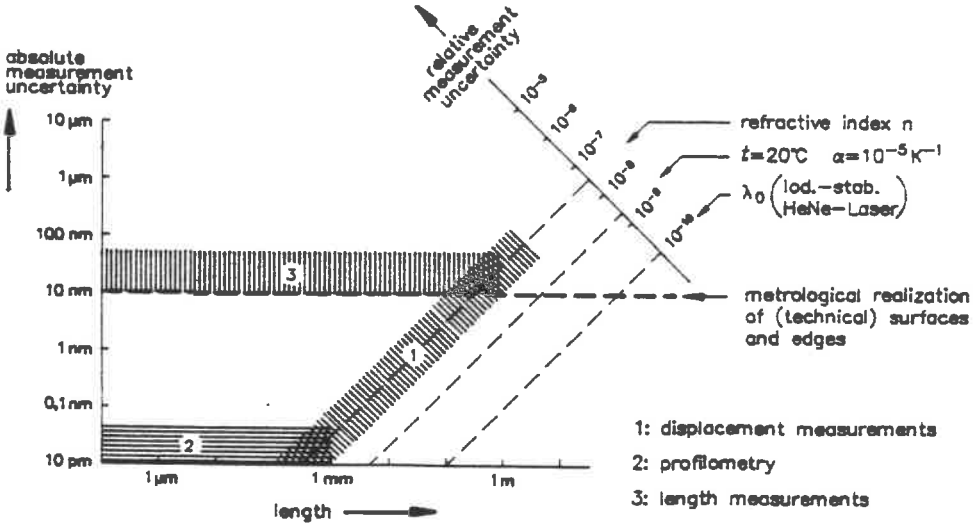


Fig.5 Limits of accuracy in dimensional metrology

SERVO-GUIDED PLANAR STAGES and OTHER X-Y STAGES FOR MICROLITHOGRAPHY : A REVIEW

Rodney Kendall
IBM Corp. Advanced Technology Center, East Fishkill, NY 12533

Extended Abstract

The development of lithography equipment in the semiconductor industry is driven primarily by the DRAM business. Approximately every three years a new generation of DRAM is bought to market, offering a four fold increase in storage over its predecessor. Each generation of DRAM requires advances in lithography tooling.

The basic building blocks of any lithography tool are the projection system, the registration and alignment system, and the X-Y stage system.

Stages fall into two broad categories ; steppers for 'step and repeat' lithography, and scanners for 'write on the fly' applications. Both types have been used with great success.

Table 1 compares stages from 13 different lithography tools. Despite the requirement that all lithography stages perform essentially the same function, it is readily apparent that there is no such thing as a 'standard' X-Y stage. Of the 13 stages described 10 use some form of fixed way and bearing combination for guidance along the X and Y axes. The remaining 3 are servo-guided stages. The servo-guided stages have no ways or guide bearings at all. Instead both guidance and positioning are provided by the drive system via the closed loop position servos. In each case the stage operates on a planar surface to constrain motion in Z and tilt, while providing freedom of motion in the X-Y plane.

At IBM a servo-guided planar stepper stage (1) has been developed to operate inside the 30 mm gap of a magnetic projection lens used for e-beam lithography. One of the design requirements for this stage was that all moving portions of the stage be completely nonmagnetic and contain no bulk metals or conductors. Guidance and positioning are achieved by the simultaneous operation of three identical servo controlled friction drives.

Configuration

Figure 1 is a schematic showing the configuration of the IBM servo-guided stage. The following descriptions refer to the EL3+ e-beam lithography system, with some results shown in figure 4.

Surface plate. The steel surface plate is a lightweighted egg-crate sandwich structure. Embedded in the top surface of the surface plate is a 254 mm dia. x 13 mm thick ferrite ceramic disc which forms the lower pole piece of the magnetic lens. The top surface, including the ferrite disc, is flat to < 5 microns.

Stage plate. The Zerodur (2) stage plate rides on the surface plate supported on 3 flat PTFE bearing pads 40 mm in diameter. Structural materials on the stage plate are either glass, ceramic or a 30% glass-filled thermoplastic. All fasteners are made from a high resistivity titanium alloy. The total weight of the stage plate is 4.2 kg. , less than 2% of which is metal.

Friction Drives. The friction drives have 300 mm of travel while allowing the drive-bar to pivot in the horizontal plane, perpendicular to the capstan axis (3). The drive bars and capstans are made from silicon nitride ceramic. The drive bars are 590 x 25 x 12 mm and weigh 0.57 kg. Capstan dia. is 12 mm.
Gauge bars. When the stage is initialized from some random position or orientation, the friction drives slowly retract such that the stage plate is sequentially banked against the two stainless steel gauge bars and the laser system reset. The 'go to gauge' function takes approx. 30 secs. and is only done once at power up.

Operation

The stage plate is kinematically located in the system by 6 constraints. The 3 stage bearing pads, the 2 drive bars in Y and the 1 drive bar in X. It can be seen therefore that controlled extension or retraction of the 3 drive bars will result in constrained motion of the stage plate in 3 degrees of freedom, namely X, Y and θz (figure 2). The normal tendency of the stage plate is to move along curvilinear paths defined by the extended length of the drive bars. In practice however the path of the stage is forced to be essentially rectilinear as the closed loop servo system continually corrects for the curvilinear path. When the stage is commanded to move along one axis the ability of the off-axis to maintain a fixed coordinate we have labeled following error. Dynamic yaw control is achieved by moving the Y drive bars differentially.

Servo System

Stage position is controlled by 3 velocity servos, 2 position servos and a yaw servo. (Figure 3).
Velocity servos. The tachometer based velocity servo is a minor servo loop within the major position servo loop.
Position servos. The position servo is provided by an HP programmable servo axis board which includes the lambda/128 (5nm/lsb) resolution extension and counter (4). When a new destination is sent to a servo axis board by the host, a series of reference positions are generated at a 4 kHz rate based on preprogrammed max. velocity, max. acceleration and max. delta acceleration parameters. The difference between the reference and actual positions, the following error, is fed into a difference equation to provide lag compensation. The reference position is fed directly to a digital feed forward section where open loop compensation for friction, inertia and slew rate is generated.

The feed forward compensation output is summed with the difference equation output and sent to a DAC which provides the analog velocity desired voltage for the velocity servo. The X DAC goes directly to the X velocity servo while the Y DAC is added to the +/- Yaw correction before going to the two Y velocity servos. The position servos have a bandwidth of 33 Hz and a phase margin of 47 Degrees.
Yaw Servo. The yaw servo is a high gain analog servo that superimposes corrections differentially with the Y velocity servos. The yaw servo has a bandwidth of 50 Hz and a phase margin of 31 degrees.

Summary

The servo-guided stage offers a performance equal to that of state of the art 'way and bearing' guided stepper stages. Design simplicity, low mass, ease of assembly, and the use of a 'standard' servo system make this type of stage a practical choice from both the technical and economic viewpoint.

MANUFACTURER AND MODEL NUMBER	APPLICATION AND OPERATION	STAGE ENVIRONMENT	FINE X-Y STAGE ON COARSE X-Y STAGE	COARSE STAGE	
				GUIDE SYSTEM AND CONFIGURATION	DRIVE SYSTEM
ASM PAS 5500	PHOTO. STPR.	AIR	NO	AIR BEARINGS, H-FRAME & PLANAR SURFACE.	LINEAR DC SERVO MOTORS
ATEQ 2564	LSR.M.M. SCNR.	AIR	NO	CROSSED ROLLER SLIDES STACKED X & Y	DC MOTORS, LEAD SCREW & PLASTIC NUT.
AT&T EBES 4	EB.M.M. SCNR.	VAC.	NO	HYDROSTATIC BRGS. ON CANTILEVERED T-STRUCTURE.	HYDRAULIC
CANON FPA2000i L	PHOTO. STPR.	AIR	NO	LINEAR ROLLER BRGS. FLAT WAYS, GUIDE BAR & SLIDER, STACKED X&Y	DC MOTORS BALL SCREWS
ETEC MEBES 3	EB.M.M. SCNR.	VAC.	NO	CYL. AIR BRGS. IN ORTHOGONAL PAIRS. EXTERNAL TO VACUUM.	DC MOTORS BALL SCREWS
GCA XLS 7000	PHOTO. STPR.	AIR	PIEZOS & FLEXURES	PTFE BRG. PADS. FLAT WAYS STACKED X & Y	DC MOTORS RACK AND PINION
HAMPSHIRE INSTRUMENTS 3500	X-RAY STPR.	He	NO	GAS(He)/VAC BRGS. STACKED PLANAR SURFS. WITH LINEAR GUIDES.	LINEAR DC SERVO MOTORS
HITACHI	EB.D.W. STPR.	VAC.	PIEZOS & FLEXURES	PTFE BRGS. WITH CROSS PLATE IN PLANAR SURFACE.	DC MOTORS BALL SCREWS
IBM EL3+	EB.D.W. STPR.	VAC.	NO	PTFE BRGS. ON PLANAR SURFACE. SERVO GUIDED.	3 PIVOTING FRICTION DRIVES.
NIKON NSR 2005	PHOTO. STPR.	AIR	NO	NEEDLE ROLLERS, VEE & FLAT WAYS STACKED X & Y	DC MOTORS LEAD SCREWS
SVGL MICRASCAN	PHOTO. SCNR.	AIR	NO	AIR BEARINGS ON PLANAR SURFACE. SERVO GUIDED.	ORTHOGONAL ELECTRO-MAGNETIC LINEAR MOTORS.
ULTRATECH 1500	PHOTO. STPR.	AIR	NO	AIR BRGS. ON T-BAR AND PLANAR SURFACE.	FRICTION DRIVES.
ULTRATECH 2000	PHOTO. STPR.	AIR	NO	AIR BEARINGS ON PLANAR SURFACE. SERVO GUIDED.	2 ORTHOGONAL PAIRS OF SAWYER MOTORS.

PHOTO	: PHOTOLITHOGRAPHY	EB.D.W.	: E-BEAM DIRECT WRITE
LSR.M.M.	: LASER MASK MAKER	STPR.	: STEPPER STAGE
EB.M.M.	: E-BEAM MASK MAKER	SCNR.	: SCANNING STAGE

TABLE 1 : COMPARISON OF SOME LITHOGRAPHY STAGES

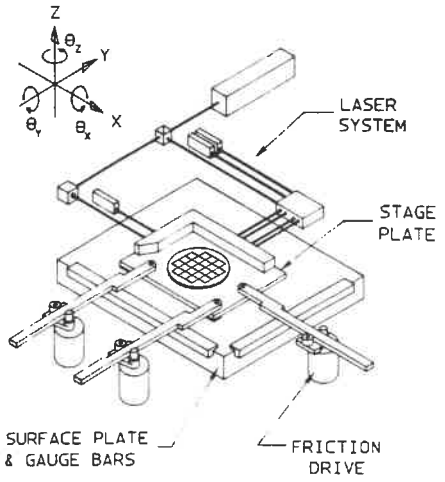


FIGURE 1. MECHANICAL CONFIGURATION OF THE SERVO GUIDED X-Y-θ STAGE.

IN EL3+ THE DUAL AXIS INTERFEROMETER AND CORRESPONDING MIRROR ARE ON THE OPPOSITE SIDE OF THE STAGE PLATE TO ALLOW EASIER ACCESS FOR LOADING AND UNLOADING WAFERS.

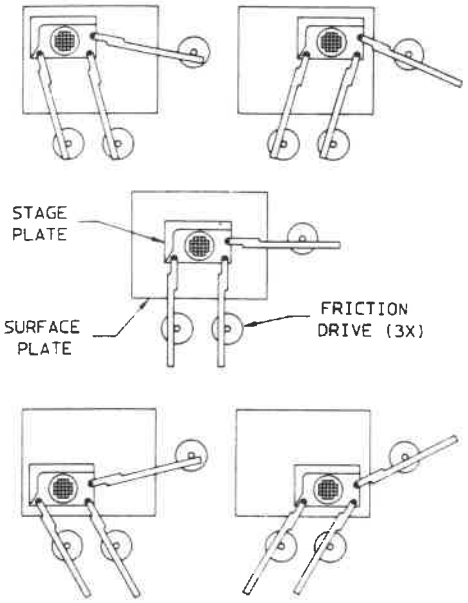


FIGURE 2. SCHEMATIC SHOWING THE MECHANICAL OPERATING PRINCIPLE

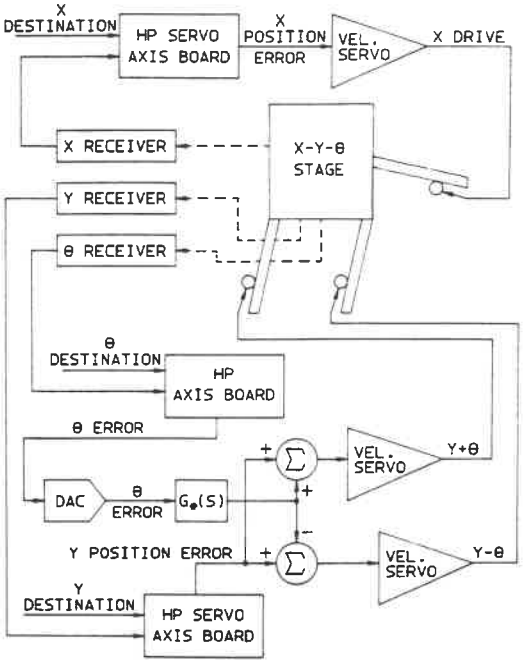


FIGURE 3
SERVO SYSTEM BLOCK DIAGRAM

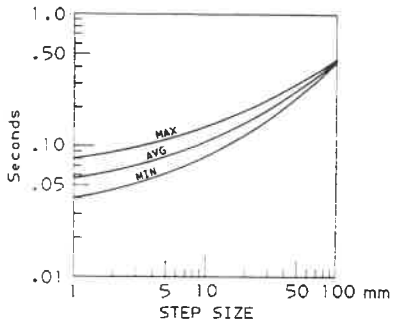


FIGURE 4a: THE STEP AND SETTLE TIME VARIES FOR ANY GIVEN STEP SIZE AS THE INTERACTION BETWEEN THE X AND Y DRIVES CHANGES WITH STAGE POSITION. IN THIS FIGURE A MOVE IS CONSIDERED TO BE COMPLETE WHEN THE STAGE IS WITHIN 1 μm OF THE DESIRED DESTINATION.

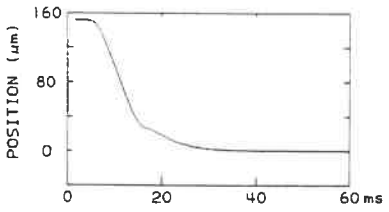


FIGURE 4b .SHOWING POSITION vs. TIME FOR A SMALL (150 μm) MOVE. POSITION ERROR AT 60 ms WAS MEASURED TO BE .04 μm. OF NOTE IS THE TOTAL LACK OF OVERSHOOT DURING DECEL. AND SETTLING.

References

(1) R.Kendall, S.Doran, E.Weissmann. EIPB'91 Conference,Seattle. To be published in J. Vac. Sci. Technol.
(2) Zerodur is a trademark of SCHOTT GLASWERKE, MAINZ, FRG.
(3) J.Reed, S.Hansen, O.Otto. J.Vac.Sci.Technol B.3 112 (1985)
(4) Hewlett Packard application note 325-10 (June 1989) and 5527A Laser Position Transducer , Designers Guide (Nov. 1986)

APPLICATION AND PERFORMANCE OF CAPACITANCE MICROMETRY AS A SUPER PRECISION TRANSFER STANDARD

S.Harb, D.G.Chetwynd, S.T.Smith
Centre for Microengineering and Metrology, University of Warwick
Coventry. CV4 7AL, UK.

Introduction

Capacitance micrometry is frequently used for high precision measurement of linear displacements, particularly at the nanometre level, Jones and Richards, 1973, Hicks et al., 1984. The accuracy to which a capacitance may be measured is limited by thermal noise of the accumulated charge and the technique for discriminating the energy stored as charge in the capacitor, and so usually on the stability of a reference device. The accuracy of a gauge depends further on achieving a sensitive and reliable geometry for converting displacement to change in capacitance. Parallel flat plate electrodes with their planes perpendicular to the axis of displacement are attractive in this respect. In practice, it is very difficult to generate a pure linear displacement free from finite parasitic tilts. If these error sources are adequately controllable, capacitance micrometry is a natural choice for transfering the absolute displacement calibration of x-ray interferometry to other precision systems. This paper explores the practicallity of so doing, concentrating particularly on the errors due to tilt.

X-ray interferometry, invented by Bonse and Hart in 1965, is now an established method for the calibration of microdisplacement transducers with sub-nanometre accuracy Schwarzenberger et al., 1989, Bowen et al., 1990. Utilising the diffraction and recombination of x-rays in a crystal lattice, systems are typically produced from a monolithic single silicon crystal. Translation of one part of the interferometer relative to its stationary base will result in the intensity of the output beam varying sinusoidally with displacement at a pitch equal to the atomic spacing of the lattice. Interpolation of this results in a known displacement to better than 30 pm, Schwarzenberger et al., 1989. The traverse is driven by an electromagnetic force transducer consisting of a small saturated permanent magnet glued to an arm on the monolith stage with a solenoid coil placed around it. Since the silicon springs are almost ideally elastic and the deflection is closely proportional to the coil current. Useful output signals are obtained only if rotation of the moving component of the interferometer relative to the stationary base is held below around 10^{-7} radians about the axes in the plane of the capacitance electrodes.

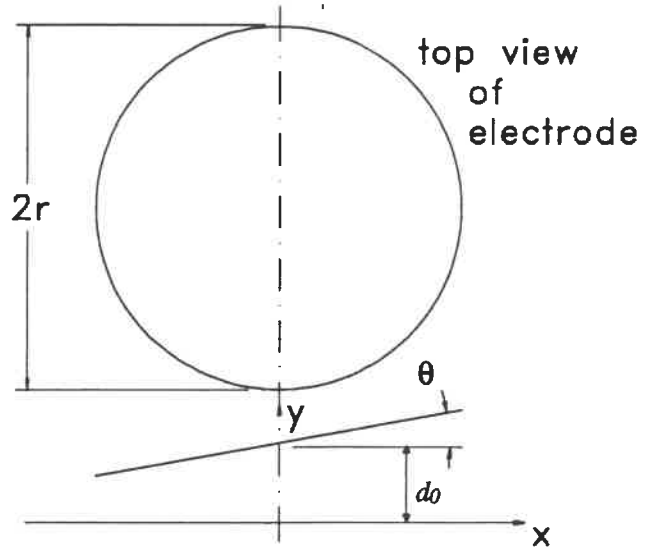


Figure 1: Mathematical model for parallel plate capacitors

The extreme sensitivity of the interferometer to parasitic twists make it an ideal technique for the evaluation of capacitance micrometers. As mentioned, any undesirable force on the moving stage of the interferometer will introduce parasitic errors that often destroy fringe contrast of the interferometer. Consequently, it is necessary to 'follow' the interferometer using non-influencing techniques. The small electrostatic forces in capacitance micrometry make it an ideal follower with the non contact electrode attached to a more robust actuator system. This is provided by a piezoelectric actuator that is feedback controlled through an internal capacitance micrometer to reduce its inherent hysteresis and thermal and temporal drifts, Hicks et al., 1984.

Theory

Consider the circular, flat plate capacitor of figure 1. The lower electrode is an infinite horizontal plane with the upper electrode separated from it by d_o at its centre and inclined at an angle θ. In plan view the tilted electrode appears as an ellipse with its major axis coincident with the tilt axis. Assuming that the gap has a homogeneous permittivity ε, there is a uniform potential over each plate, and fringing fields can be ignored, the capacitance, C, is obtained by integrating over elemental parallel plates of width dx and separation y to give

$$C = 2\varepsilon \int_{-r\cos(\theta)}^{r\cos(\theta)} \frac{\left(r^2 - \left(\frac{x}{\cos(\theta)}\right)^2\right)^{1/2}}{d_o + x\sin(\theta)}\,dx = \frac{2\varepsilon r}{d_o} \int_{-r\cos(\theta)}^{r\cos(\theta)} \frac{\left(1 - \left(\frac{x}{r\cos(\theta)}\right)^2\right)^{1/2}}{1 + x\sin(\theta)/d_o}\,dx \qquad (1)$$

This can be integrated and rearranged to give an explicit expression for the capacitance C given by

$$C = \frac{2\pi\varepsilon r^2\cos(\theta)}{d_o}\left[\frac{1 - (1 - k^2)^{1/2}}{k^2}\right] \quad \text{where } k = \frac{r\cos(\theta)\sin(\theta)}{d_o} = \frac{r\sin(2\theta)}{2d_o} \qquad (2)$$

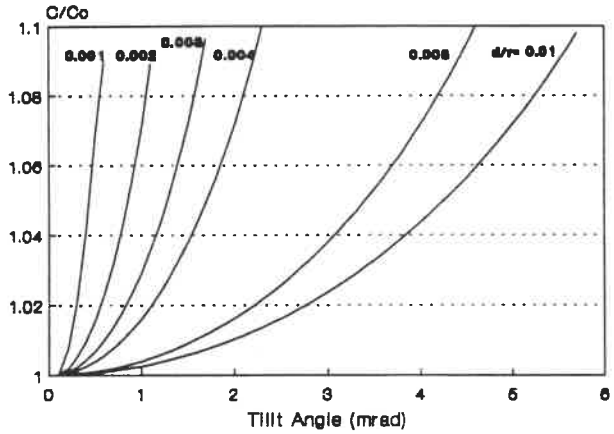


Figure 2: Normalised ratio of capacitance value as a function of tilt for the radius to diameter ratios d/r ranging from 0.001 to 0.01

The range of $|k|$ lies between zero (parallel plates) and unity, when the two electrodes touch and the original model is no longer valid. Equation 2 converges to the true parallel plate formula as θ tends to zero. For computational purposes it is convenient to expand it using binomial series . Ignoring terms of the order k^8 and normalising to the capacitance value for parallel plates, C_o, equation 2 becomes

$$\frac{C}{C_o} = \cos(\theta)\left[1 + \frac{k^2}{4} + \frac{k^4}{8} + \frac{5k^6}{64} + \ldots\right] \qquad (3)$$

This has been plotted in figure 2 for varying ratios of the radius of the electrodes to average separation over a range typical for high sensitivity gauges. Error terms up to the twelfth order are included in this computation. As the electrode separation is reduced there is a marked increase in sensitivity to angular changes which will be discussed later.

Experimentation

The complete mechanical instrumentation for these experiments is shown in figure 3. It includes a Queensgate Instruments AX100 Digital Piezoelectric Transducer (DPT) rigidly attached to an aluminium plate that is kinematically supported by three threaded legs resting in grooves machined into the x-ray interferometer monolith.

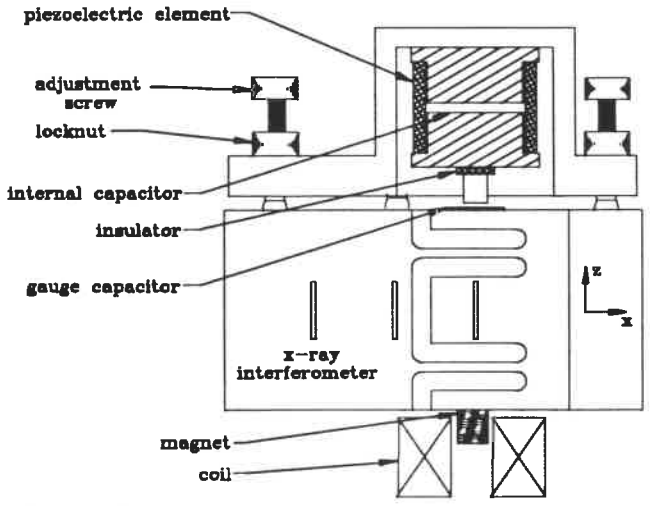


Figure 3: Experimental setup for the calibration of a capacitance micrometer transfer standard calibration

These supports enable linear adjustment in the z-axis and angular adjustments about the x and y axes. The lower face of the DPT carries a 6 mm diameter flat faced brass electrode. This forms a parallel plate capacitor with a secondary electrode cemented to the top of the monolith stage, which is gold plated to provide a conduction path. The nominal design value for this external capacitor was 10 pF and its clearance 10 μm giving an electrode separation to radius ratio of 1:300 or 0.0033. It was possible to obtain a repeatable gap separation to within approximately 100 nm and angular adjustment (although not parallelism) to within 10 arc seconds (0.05mrads). The drive to the monolith used a 6 mm diameter, 5 mm long Co-Sm magnet and a 15 mm diameter coil of about 200 turns. This provided a sensitivity of about 1 nm/mA and displacements were restricted to a

maximum of 20 nm to minimize heating near to the sensors. The complete assembly was mounted on a large concrete block supported by air spring vibration isolators and housed in a temperature controlled laboratory.

A cascade control system with two modes of operation is used. The DPT is driven by an autonomous proportional plus integral controller that compares an input position demand to the gap of its internal capacitive sensor to provide the drive signal for the piezo element. The controller is switchable to use either an internal or external reference signal. The external gauge capacitor is monitored and conditioned by a special capacitance bridge circuit (provided by Queensgate Instruments) which is used to provide the position demand signal to the DPT. Both capacitive gauge signals are monitored by a computer, which also directly drives the coil of the monolith traverse. The two operational modes differ in the manner of the signal interconnections. Brief, if somewhat imprecise, labels for these will be used here: open and closed loop modes. In open loop mode, the internal piezoelectric gauge is maintained at a constant value, which, in turn, maintains the length of the actuator. Motion of the interferometer is directly monitored by the external sensor which can then be calibrated. This information is used to assess the long term stability of the instrumentation. The closed loop mode uses the external sensor signal as the error input to the DPT so that it follows the x-ray interferometer. In this mode, the displacement of the actuator and interferometer can be measured from the internal sensor signal, as well as from the coil current or, for absolute calibration, the x-ray intensity fringes. The effective discrimination in measurements is approximately 30 pm and 3 pm for the DPT and external gauge respectively. Temperature varied by no more than 0.1 K during any one calibration experiment.

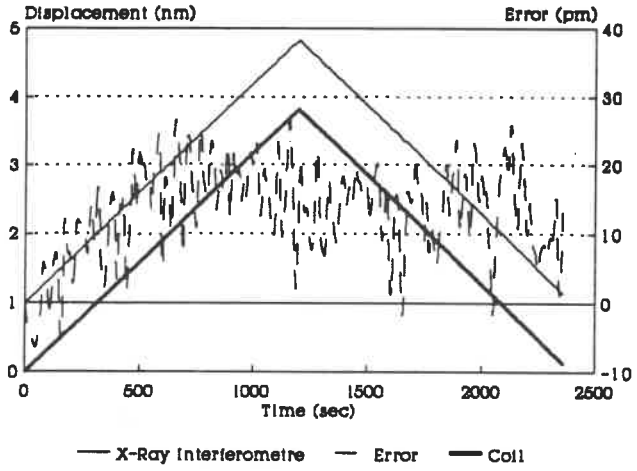


Figure 4: Monolith stage displacement calibrated by three phase x-ray interferometry for a coil demand of 3.8 nm.

Results

Calibration of the electro-magnetic drive of the monolith stage against the x-ray interferometer indicated that it had extremely good linearity, accuracy and repeatibility. Figure 4 shows a typical output from a calibration over 3.8 nm traverse. The output signal from the forward diffracted x-ray beam was analysed using a three phase stepping technique. The reconstructed displacement shows errors of up to 30 pm. The drive is closely similar to that used by Schwarzenberger et al., 1989, and the reduced accuracy may be attributed to heavily assymmetric phase steps and short (3 s) count times. The characteristic oscillation at exactly half the fringe pitch in the reconstruction is indicative of a systematic error from these sources. The underlying trend of the error curve

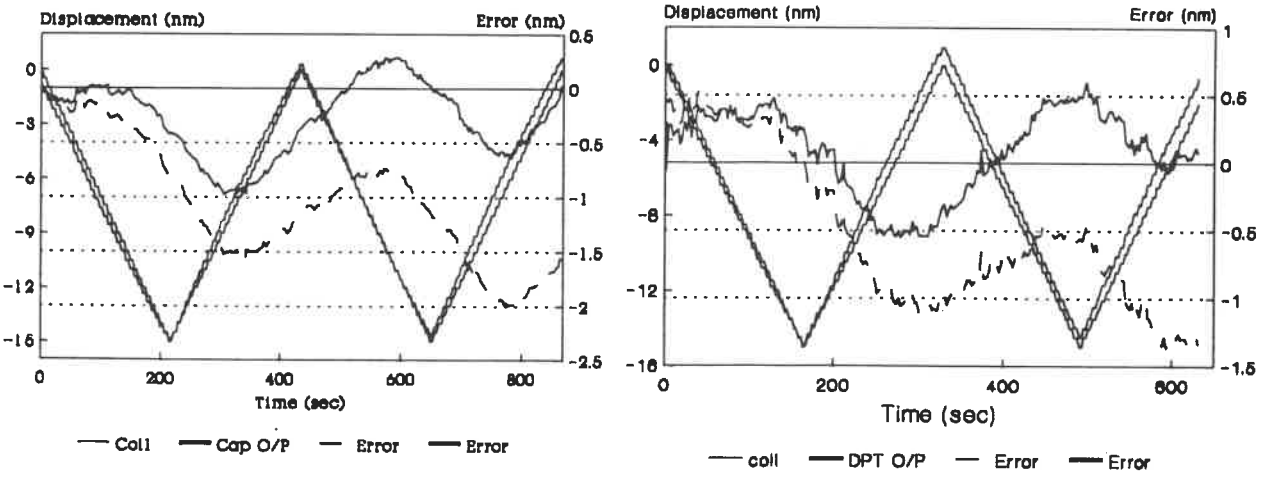


Figure 5: Displacement characteristics of the capacitance gauges as a function of the stage displacement; a) external sensor in open loop mode and b) internal sensor in closed loop mode.

indicates that the coil drive is good to better than 20 pm without explicit use of x-ray interferometry so that faster calibrations of the capacitors and follower can be undertaken.

Calibration of the external capacitor gauge was performed in the open loop configuration by commanding the coil drive in equal steps with a delay of 1 second between each. At each position the coil input and gauge signal were measured and recorded. The coil was driven in a triangular wave to provide bi-directional motion so that hysteresis and backlash might be detected. For each test the nominal gain of the sensor was estimated from the data by using a least squares fit to estimate the gradient and offset coefficients, a and b, of an assumed linear relationship between the gauge and the corresponding coil position. These values were then used to calculate the error. An almost identical procedure was used to calibrate the internal capacitor of the DPT with the system operating in closed loop mode. A typical calibration result is shown in figure 5. Figure 5(a) shows the output of the external capacitor gauge in the open loop mode alongside the coil demand, while in Figure 5(b) shows the output of the DPT with the coil demand when operating in closed loop control. The difference between coil demand and gauge output is plotted as 'error'. The drift in both instances can be attributed mainly to slow thermal changes which can be compensated as shown.

The error between capacitance output value and displacement appears to be systematic with a phase lead in this particular instance. Numerous experiments reveal that, not only will the magnitude of this systematic component vary, but the phase can jump from a lead to a lag after each time the test rig is broken down and reassembled. Previous error analysis by Harb et al., 1991, has shown that error sources other than tilt are likely to be insignificant at this level or do not explain the observed phenomena. Differentiating equation 3 allows the prediction of tilt sensitivity at any given offset angle, θ. For the dimensions of external gauge used here the apparent displacement is about $10^8\theta\delta\theta$ nm. For example with an offset of 1 mrad, an error of 0.2 nm would be caused by a parasitic twist of around 2 μrad during the 15 nm traverse.

Discussion

It has been shown that controlled piezoelectric actuators with in-built capacitive sensing are capable of defining displacements to a precision of at least 0.1 nm. They can provide a high-stiffness follower for reproducing the motion of the interferometer to allow the calibration of other devices with a traceability to, at worst, 0.4 nm (of the order of the atomic spacing of many crystals). The capacitance micrometer can be used as a transfer standard for small displacements with a traceability of no more than this value in the worst case and with a present bandwidth of 1 kHz.

The effects of tilt on the accuracy of parallel plate capacitance gauging have also been investigated. It has been shown that, as might be expected, the sensitivity to angular tilt errors varies strongly with electrode separation. However, compensation may be applied to this systematic error and using the closed form solution for small tilts the magnitude of this correction in terms of an angular error may then be estimated.

Making conservative assumptions about the degree of systematic component in the errors observed and noting the contribution of the calibration uncertainty of the evaluation system itself, allows prediction of what should be achievable without the need for new technological approaches. Both as a stand alone actuator and as a follower operating at a few Hz bandwidth the DPT system should provide discrimination in the 20 to 50 pm range. As a follower this would also be the limit on traceability. The capacitance micrometer should be capable of better traceable precision with a reasonable bandwidth, presently up to 1 kHz. Transfer standards involving physical repositioning are likely to incur a slight loss of precision. These figures still represent an upper bound on the performance of the systems; it is quite likely that significant improvements on them will be demonstrated within the next few years.

References

Bonse U. and Hart M., 1965, "An x-ray interferometer", *App. Phys. Letts.*, Vol. 6, No. 8, pp. 155-156.

Bowen D.K, Chetwynd D.G. and Schwarzenberger D.R., 1990, 'Sub-nanometre displacements calibration using x-ray interferometry', *Meas. Sci. Technol.*, Vol. 1, 107-119.

Harb S., Smith S.T., and Chetwynd D.G., 1991, 'Sub-nanometre behaviour of a capacitive feedback, piezoelectric displacement actuator (monitored by x-ray interferometry)', *Rev. Sci. Instrum.*, Submitted.

Hicks T. R., Reay N. K. and Atherton P. D., 1984, "The application of capacitance micrometry to the control of Fabry-Perot etalons", *J. Phys. E: Sci. Instrum.*, Vol. 17, pp.49-55.

Jones R.V. and Richards J.C.S., 1973, "The design and some applications of sensitive capacitance micrometers", *J. Phys. E.*, Vol. 6, pp. 589-600.

Schwarzenberger D.R., Chetwynd D.G. and Bowen D.K., 1989, "Phase measurement x-ray interferometry", *J. X-ray Sci. Technol.*, **1**, 134-142.

Advanced Angle Metrology at the National Institute of Standards and Technology

W. Tyler Estler, Yun Hsia Queen, and David Gilsinn

Precision Engineering Division
National Institute of Standards and Technology
Gaithersburg, MD 20899

and

Dan (Pie) Pieczulewski
A.G. Davis Gage & Engineering Co.
21435 Dequindre, Hazel Park, MI 48030

Until recently, the basic techniques for calibrating indexing tables, precision polygon mirrors, and rotary encoders at NIST and other metrology laboratories had not changed greatly in more than thirty years. A series of round-robin intercomparisons involving eight national laboratories demonstrated variability of approximately $\pm$ 0.2 arc-sec ($\pm$ 1 μrad) in the measured face-to-face errors of two 12-sided polygon mirrors. This is a typical level of repeatability observed at NIST for precision artifacts with long measurement histories.

In studies supported by the Calibration Coordination Group of the Department of Defense and, more recently, by the Strategic Defense Initiative Organization, we identified a number of factors which conspire to limit the achievable accuracy of angular measurements and calibrations. Principal among these factors are (1) thermal drift associated with manual procedures, (2) refraction and air turbulence effects on small-angle measurements, (3) seismic and acoustic disturbances, and (4) optical errors in autocollimators and polygon mirrors. We have recently installed a new automated angle calibration facility at NIST designed to minimize measurement errors and achieve a factor of ten improvement in full-circle angle metrology.

Called the Advanced Automated Master Angle Calibration System (AAMACS),

the new facility has three major components: (a) a very high resolution indexing table system, (b) a thermally stable vibration isolated custom optical table, and (c) a high resolution laser interferometric small angle measuring system. The heart of the apparatus is a set of three stacked independently driven serrated tooth indexing tables arranged to rotate concentrically around a common vertical axis. This stack is topped by a precision air bearing servo table with integral motor/optical encoder package. The stacked indexing tables form a rotary analog of a differential screw. The tooth numbers of these tables ($832 = 13 \cdot 2^6$, $729 = 3^6$, and $625 = 5^4$ respectively) are relatively prime integers, so that the number of discrete angular positions of the topmost table is $832 \cdot 729 \cdot 625 = 379{,}080{,}000$. Conversely, the least angular step between nearest neighbor positions of the topmost table is

$$\begin{aligned} \Delta\theta_{min} &= 2\pi/(379{,}080{,}000) \text{ rad} \\ &= 17 \text{ nrad } (0.003 \text{ arc-sec}). \end{aligned}$$

The indexing tables are capable of simultaneous motion and continuous rotation in either direction. They are driven by capstan systems through closed-loop DC servo motors using resolver feedback to locate the chosen tooth addresses. The capstan drives are disengaged when the tables are locked, and the lock forces are provided by Belleville springs so as to be independent of supply pressure to the air bearings.

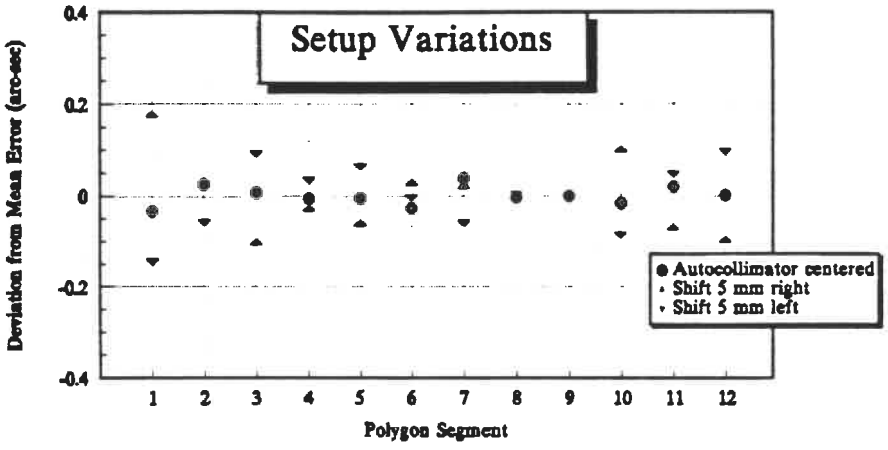
A system controller provides for completely automatic operation, so that repeatable angular positioning errors can be measured, and a system error map created, with no operator intervention. The calibration algorithm is a modified version of a well-known technique used to compare and simultaneously calibrate two nominally identical indexing tables. The basic idea is to choose one of the three indexing tables, whose errors are to be determined, and to compare it with a second "table" synthesized from coordinated motions of the other two. Each of the three tables may be self-calibrated in turn. The error mapping process has not been completed.

Small angular errors are measured using a two-frequency heterodyne HeNe laser system using polarization optics and a difference frequency of 250 kHz. The phase difference between the reference and signal sine waves is measured with a commercial precision phasemeter. As currently configured, one degree of electrical phase at 250 kHz corresponds to approximately 0.02 arc-sec (100 nrad) of rotation of the target mirror. Electrical phase measuring accuracy at 250 kHz is near $\pm 0.2°$, yielding small-angle accuracy of $\pm 0.004°$ in an interferometer system free from noise and non-linearity errors. The current interferometer is noise-limited at a level of approximately ± 0.02 arc-sec RMS, and has not been corrected for polarization mixing errors. With this level of primarily turbulence noise it is

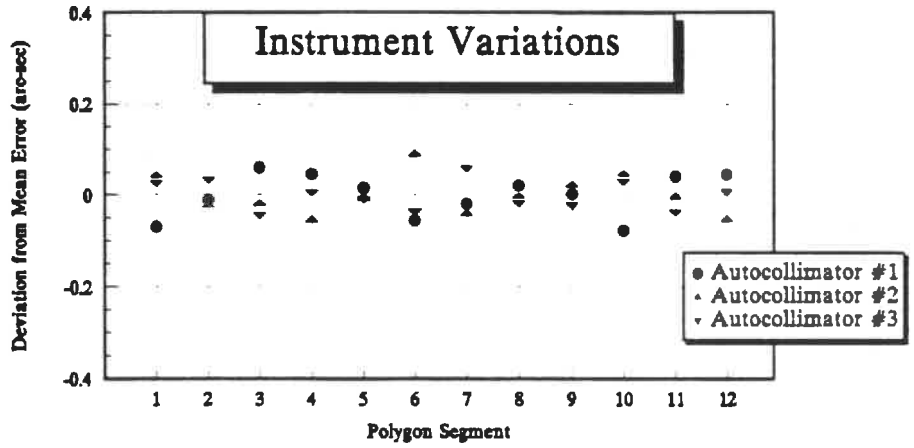
interesting to note that we have been unable to measure any lack of position repeatability in the AAMACS indexing table stack. The servo table, by way of comparison, has a 3σ repeatability of approximately ± 0.2 arc-sec (± 1 μrad).

We have recently completed a study of the systematic errors associated with the calibration of precision polygon mirrors used as angle transfer standards. A 12-sided (30°) NIST polygon with a long measurement history was repeatedly measured with the AAMACS system, using a variety of setup geometries and high-resolution autocollimators. By using multiple orientations and the constraint of circle closure (i.e. angle errors must sum to zero), polygon errors were measured independent of the AAMACS systematic errors. For any given setup, the 3σ measurement process repeatability was near ± 0.02 arc-sec (± 100 nrad). The results of these experiments are shown in Figure 1. The clear conclusion from this data is that polygon/autocollimator systems are not satisfactory for angular measurements where absolute accuracy of better than ± 0.2 arc-sec (± 1 μrad) is required.

NIST 12-Sided Polygon S/N 325



(a)



(b)

Figure 1

Results of high-accuracy polygon calibrations are sensitive to the details of the setup geometry and variations in the optical systems of different autocollimators. Shown above in (a) are the mean deviation errors resulting from 5 mm lateral shifts (left and right) of a precision autocollimator nominally centered on the polygon faces (autocollimator aperture = 50 mm, polygon mirrors rectangular 20mm x 25 mm). In (b) are shown the corresponding deviations observed when different autocollimators are used to measure the same polygon. The mirrored faces of the polygon have peak-to-valley figure errors in the range $\lambda/10$ - $\lambda/20$, with systematic roll-off near the edges. Each data point is the average of 10 measurements. The 3σ measurement process uncertainty for each calibration run using the AAMACS system is typically $\pm$ 0.02 arc-sec ($\pm$ 100 nrad).

SCALES vs INTERFEROMETERS: A COMPARISON OF METROLOGY SYSTEMS

Geert Wyntjes, Donald Mitchell and Michael Hercher
OPTRA Inc. 66 Cherry Hill Drive
Beverly, MA 01915 ✆508-921-2100

Summary

Sensors using interferometric readout of transverse scales have recently achieved resolutions comparable to those achieved with conventional axial Michelson interferometers using 2-frequency lasers. For scales, the metric is the scale spacing while for an axial interferometer the metric is the wavelength of the laser in air. To the extent that scales can be made less sensitive to atmospheric variations than the wavelength of a stabilized laser in air, the scale offers the potential for improved precision in situations where measurements are made in a real, variable atmosphere.

Although an interferometric readout device is employed with a scale metrology system, the optical path difference is nominally zero and does not change during the course of a measurement: to first order, the measurement is independent of wavelength and wavelength stability of the source is not an issue. The accuracy of a scale system depends on the the precision with which the scale has been made, the accuracy of its calibration, and the accuracy with which the interpolation is carried out. Somewhat ironically, the best scales are made in a stable environment using a 2-frequency laser interferometer to control the line spacing.

We will review and compare convenience, cost and sources of error in both types of metrology systems, and will present a variety of comparative data.

Basic principles

2-FREQUENCY LASER INTERFEROMETER: the 2-frequency laser interferometer is a Michelson interferometer which is illuminated by the collimated beam from a 2-frequency stabilized HeNe laser. The two frequency components are orthogonally polarized and can be separated by a polarizing beamsplitter. One frequency goes into the stationary reference arm of the interferometer and the other goes into the other arm of the interferometer—which contains a movable mirror whose displacement is what is measured. The interferometer mirrors are generally corner-cubes to make them insensitive to small tilts. The HeNe laser frequency can be stabilized to significantly better than 1 part in 10^8: errors in frequency stabilization are not a significant factor in measurement accuracy with this type of laser interferometer.[1] It is the refractive index of air, which depends on pressure, temperature, relative humidity and CO_2 concentration, which plays the dominant role in limiting the accuracy of measurements made with a 2-frequency laser interferometer.[2] For very high resolution measurements, cross-talk (i.e. modulation of the laser intensity at the difference frequency) may also play a role in limiting accuracy.

---

[1] Mark A. Zumberge, "Frequency stability of a Zeeman-stabilized laser," Appl. Opt.**24**, 1902-1904, 1985.

[2] W. Tyler Estler, "High-accuracy displacement interferometry in air," Appl.Opt.**24**, 808-815, 1985.

With a 2-frequency laser interferometer there is a reference signal derived from the laser which is modulated at the difference frequency which is typically in the neighborhood of 1.6 MHz, and there is a signal frequency which is at the reference frequency when the mirror is stationary, and which sees ±1 additional cycle of phase for each 1/2-wavelength displacement of the mirror. Displacements smaller than 1/2-wavelength can be measured by measuring the phase difference between the signal and reference signals. Because these signals are both modulated at about 1.6 MHz, their phase difference can generally be measured in the time domain, ensuring good accuracy in the phase measurement (unless cross-talk is present).

SCALE-BASED MEASUREMENT: In a scale-based measurement, overall accuracy resides in the scale and resolution and precision are determined by the readout sensor. The only environmental factor affecting accuracy is temperature, which determines the thermal expansion of the scale and of the mechanical structures linking the scale and the readout sensor. The type of scales with which we are concerned are conventional linear encoders which can be inexpensively reproduced and which have spatial period on the order of 0.0005-inch to 0.001-inch (12.5-micron to 25-micron). Resolution and precision are thus determined by the degree and precision with which we can interpolate between successive rulings on the scale. Using carefully assembled moiré-type readout sensors some manufacturers have been able to reliably interpolate with a resolution as good as 1/512-cycle.[3] Using a different type of sensor head we have been able to achieve resolutions of 1/3072-cycle at a 10 MHz update rate, and still higher resolutions at reduced measurement bandwidths.[4] Combining this type of readout sensor with a 1250 line/inch scale gives a resolution of 0.26-microinch (6.6 nm) at the full 10 MHz bandwidth—allowing on-the-fly full-resolution measurements at velocities up to 2.6 inches/sec; measurements can be *tracked* at velocities of more than 100 mph (160 km/hr). This latter type of scale technology can clearly match or exceed the resolution of the 2-frequency laser interferometer: differentiations between the two measurement approaches must be based on other considerations—such as achievable accuracy and precision in real measurement environments, cost, ease of installation, maximum velocity, and susceptibility to sources of error external to the metrology system (e.g. Abbe errors).

Performance issues

RESOLUTION: as discussed above, both laser interferometers and scales can achieve resolutions at or close to the nanometer level. Moreover, both approaches offer sufficient signal-to-noise to appear capable of still higher resolution. We do not feel that the two approaches can be meaningfully differentiated on the basis of resolution.

---

[3] The Itek division of BEI makes encoder readout sensors which generate sinusoidal signals in quadrature with sufficient fidelity to allow interpolation to 1/512-cycle using a scale with 1000 line pairs/inch (Donald Breslow, private communication).

[4] Michael Hercher & Geert Wyntjes, "Fine Measurements from Coarse Scales," Proc.5th Annual Conference on Precision Engineering, Monterey, CA, 1989.

PRECISION:

Laser interferometer: in a vacuum the 2-frequency laser interferometer should achieve a precision (repeatability) comparable to its resolution—which is unexcelled except by interferometers operating at shorter wavelengths. In an air environment, the laser interferometer's precision is limited by the uncertainty in the refractive index of air along the entire optical path within the interferometer. Estler[1] has addressed this issue in some detail. The table below is taken from his paper.

Parameter	*Nominal value*	*Change for which $\Delta n = \pm 1$ part in 10^7*
Pressure	760mm Hg	+0.28mm Hg
Temperature	20.0°C	-0.1°C
Humidity	40%	-10%
CO_2 conc.	340ppm	+670ppm

Clearly temperature and pressure are the major culprits: in many instances, particularly where there is a long dead path (i.e. air path in the interferometer within which no measurements are made), temperature is measured at many points along the air path in order to achieve high measurement precision. Temperature and pressure corrections may be made manually, or automatically (using an electronic "weather station" which is coupled to the data processor). Where extreme precision is needed, it is even necessary to take into account pressure variations with height if the optical path is not horizontal! Another factor to be considered is the thermal expansion of the structure which maintains the relative positions of the interferometer cube and the moving mirror. Local air turbulence and even air temperature changes due to the presence of a human observer are readily observable and contribute to measurement imprecision.

Scale: with a scale, the only significant sources of imprecision are due to uncompensated (or uncorrected) thermal expansion in the scale or in the structure which positions the readout head relative to the scale. Scales may be made of metal, which offer the advantage of uniform temperature due to high thermal conductivity, or glass. Proper selection of the material can assure thermal expansion coefficients on the order of 10^{-6}/°C or less (Invar, Zerodur, ultrasil etc.). Moreover, since the measurement of scale temperature can be accomplished with more precision than that of air temperature, it should be possible to make the relevant temperature measurements with greater precision than can be realized with a laser interferometer. Using a scale with a thermal expansion coefficient of 10^{-6}/°C, a temperature change of 0.1°C would produce a change in a measured displacement of 1 part in 10^7. This is the same as for the interferometer, except that it should be possible to make a more precise measurement of the scale temperature than of the mean air temperature. Since laser interferometer measurements in air are vulnerable to pressure variations, and are somewhat more vulnerable to temperature variations than are scale measurements, the case can be made that *in air* scale measurements can be more precise than laser measurements.

ACCURACY:

Laser interferometer: again, in a vacuum (or in a known and highly stable air environment), the laser offers very high accuracy (comparable to the resolution) when measurements are made with care. However, in air, accuracy suffers in exactly the same ways that precision

suffers—except that accuracy requires quantitative knowledge of temperature and pressure, whereas precision requires only that temperature and pressure remain stable during the series of measurements.

Scale: high precision is converted to high accuracy by way of calibration. Scale measurements are only as accurate as the accuracy with which the locations of the scale markings are known. Scales can be made very accurately or, equivalently, calibrated very accurately by the simple expedient of either making them under laser interferometric control in a stable environment (or even vacuum), or calibrating them against a laser interferometer under the same conditions. The calibration process is somewhat simplified by the fact that most scale readout schemes involve spatial averaging over a scale distance on the order of 1 mm (so that the correction table doesn't need more than 1 or 2 entries per mm at most). Since there has not in the past been any pressing incentive to make scales with very high accuracy (better than ± a few microns), it is not yet known how accurately scales can be made. Experience with ruled diffraction gratings suggest that very high accuracies can be achieved. Perhaps more important, from a manufacturing point of view, is the very high fidelity of replicated gratings made from ruled masters. The bottom line here is that scales should be able to achieve the same level of accuracy, through either a manufacturing or calibration procedure, as laser interferometers can achieve is a controlled environment. The calibration process is conceptually cumbersome, but in fact (through the use of inexpensive PROMs) quite transparent to the user. Another issue is linearization on the interpolation process: this will be addressed in some detail in the full paper. Suffice it to say that we have been able to achieve precision at the sub-microinch level using scales with 1250 lines/inch. A separate small PROM, with a dynamic range of 1-cycle, is used to linearize the interpolation.

SPEED: The maximum speed at which either system can track is simply equal to the measurement update rate divided by the displacement corresponding to 1/2-cycle. For equal update rates, the scale handles higher speeds due to the larger value of the displacement per cycle (≈ 15 microns vs. 0.63 microns). The laser interferometer has an additional velocity limitation in one direction: if the Doppler-shift in the reflected laser beam approaches the heterodyne beat frequency, then the measurement frequency approaches zero and the measurement fails because of the AC-coupling used for data processing. The speed at which this occurs for a beat frequency of 1.6 MHz is about 50 cm/sec. Our scale measurements, using a 10 MHz update rate, are theoretically capable of tracking at a velocity greater than 2500 cm/sec: we haven't been able to move anything fast enough to lose tracking.

INSTALLATION, COST & OTHER ISSUES: laser interferometers are unexcelled for measurements in highly controlled environments, but scales with high-resolution readout heads offer equal or better performance in ambient air and can be significantly less expensive. Laser diode-based readout sensors of the type we have been working with consume very little power (< 100 mwatt; possibly as low as 10 mwatt average power), are very compact, and can be designed to contain their data-processing electronics on ASIC's built into the head (our current electronics, including all power supplies, are all built onto a single plug-in card for a PC). Installation of the scale metrology system is essentially the same as for a conventional linear encoder, except that the readout head is displaced by 5 ± 0.5 mm from the scale. For XY-stage metrology systems, we have designed a planar (XY) scale with a single readout head located directly below the point of measurement: this very greatly reduces Abbe errors and leads to a very compact and potentially fast precision stage.

Scanning Probe Microscopes as Co-ordinate Measuring Machines

Margaret Stedman
National Physical Laboratory,
Teddington, Middlesex TW11 0LW, UK.

Scanning probe microscopes (SPMs) form a family of instruments in which images are built up by sequential scanning in a raster pattern, rather than simultaneously by a focusing element. In recent times a large variety of probes with high lateral as well as vertical resolution have been devised and applied in SPMs. The scanning tunnelling microscope (STM) and the atomic or scanning force microscope (AFM or SFM) are probably the most well known of the SPMs that have been developed over the last decade. These sensitive probes combined with scanning technology and advanced computer graphics have produced images with great visual impact which seem to validate the use of the term *microscope*. However, the images are constructed from dimensional data in three co-ordinates. In reality, SPMs are small co-ordinate measuring machines - CMMs in miniature. The aim of this paper is to examine their roles in dimensional metrology through an exploration of the analogy between SPMs and CMMs.

In a CMM the three dimensional motion of the probe relative to the workpiece is normally realised by combining three orthogonal linear slideways. This may be done in a variety of machine configurations. Analysis of the possible errors of carriage motion easily leads to the concept of 21 independent sources of geometrical error which combine to generate the position error of the probe relative to the workpiece being measured. An early approach to the calibration of CMMs was based on the attempt to measure all of the 21 errors of motion, which could then be combined to estimate the error of the probe co-ordinates at any position. This daunting and impractical approach has now been largely superceded by the parametric approach in which a stiff and stable artefact containing an array of measurable points, is repeatedly measured in a range of orientations and positions within the working volume of the CMM. Typical calibration artefacts would be a ball plate or a step gauge; in principle, even a simple end gauge could be used. The data obtained is capable of analysis to yield an error map of the working volume. Similar approaches to the calibration of two axis measuring machines can be made.

In an SPM the resolution and range of motion is unsuited to slideway technology, and actuation is normally obtained by distortion or flexure of an integrated or even monolithic structure such as a PZT scanner tube. The characteristics of PZT actuators (for which the drive voltage is often taken as a measure of displacement) directly affect the quality of data produced. Non-linearity, hysteresis and scan-rate dependence are

contra-indications for good metrology. The integration of the actuation and metrology functions in an SPM scanner confers the advantages of high stiffness and resonant frequency - features which may be crucial in a device achieving atomic resolution. This contrasts with CMMs in which separate displacement transducers, such as metrological gratings, are always provided. A few special SPMs with on-board metrology have been built, using for example capacitance, optical interference, optical shadowing and x-ray moire interference as position transducers. Such devices improve metrological quality, but do not necessarily obviate the need for calibration, and have not yet been included in any commercial SPMs.

Most SPMs are used as profilometers, ie the z range of motion is much less than the x and y ranges. In these circumstances the SPM might be said to be a 2.5D machine, in contrast to the 3D CMM. The effective number of independent errors of motion will be less than the 21 of a CMM. Even so, their direct calibration is still an impracticable task, whereas an analogue of the parametric CMM calibration is conceivable. Suitable artefacts exist, or can be devised. For example, a grating is an analogue of the step gauge used in CMM calibration. An array of pyramidal hills or pits with (111) faces on a silicon chip is analogous to the ball or hole plates used with CMMs. The essence is that probing the pyramidal faces, which are well defined geometrically, allows the position of the apex to be estimated, just as probing a spherical surface allows the position of its centre to be determined. Such artefacts can be made by silicon micromachining methods. However, as the scale is reduced from micrometres to nanometres, it will become increasingly difficult to achieve the necessary resolution even by electron beam writing. At the atomic resolution level, it seems inevitable that crystal lattice spacings will provide the route to calibration. However, the phenomenon of surface re-arrangement imposes a need for caution, since the atomic patterns at the surface will generally have a cell size different from, though related to, the crystal lattice parameters determined by x-ray diffraction. In principle, these artefacts could be used to make an error map of the working volume of the SPM, just as can be done with ball plates for a CMM. In practice, it will probably be more appropriate to calibrate by measuring an artefact (such as a grating or an array) at a standard position within the working volume and at a standard scan rate.

A contrast between SPMs and CMMs is that the former are used under the dynamic conditions of continuous scanning, whereas the latter are used statically. This has implications for the appropriate form of calibration artefact. Thus, for a grating artefact a sinusoidal groove form has the advatage of Fourier purity. The instrument is characterised or calibrated at a particular amplitude (A) and wavelength (W), without being affected by its characteristics at other amplitudes and wavelengths. It follows that SPMs with a wide dynamic range should be calibrated with several sinusoidal gratings of suitable AW values distributed within the AW map of the instrument.

In all probe metrology, whether it be with an SPM or a CMM, the effects of the interaction of the probe with the surface must be considered. In CMMs ball-ended probes are used, and the removal of convolution with the surface geometry is relatively simple. For SPMs considerations of stiffness and dynamic performance preclude the use of a simple analogue of the CMM probe, although a spherically-tipped cone may approximate to the shape of the best tips. Tip-surface interactions may be classified as weak or strong. The best metrolgy (and calibration) will be achieved with a weak interaction. Similarity between the probe and surface topography with regard to values of slopes and curvatures will imply a strong interaction, and the uncertainty associated with the measurement or calibration will be increased. Strong interactions are indicated when instruments are used near the edges of the AW maps.

In summary, exploration of the analogy between SPMs and CMMs highlights a range of issues in the validation of SPM metrology, and helps in developing artefacts and strategies for the traceable calibration of SPMs.

THE CHARACTERIZATION OF ANALOG PROBES FOR USE ON COORDINATE MEASURING MACHINES

Howard H. Harary
National Institute of Standards and Technology
Gaithersburg, MD USA

and

Didier Clermont, Claude Lavigne, Patrice Chollet
and André Clement
Institut Supérieur des Materiaux et de la Construction Mécanique
St. Ouen, France

In many ways, analog (proportional) probes are really miniature coordinate measuring machines (CMM). Analog probes can have scale, squareness, straightness, and other geometric errors. As long as these errors are repeatable, an error correction model can be applied to improve the probe's performance. The careful characterization of analog probes at the 0.1 μm level is worthwhile because they are often the sensor of choice on high accuracy CMM's, in some cases limiting CMM measurement accuracy. Our goal was to measure the accuracy of two types of analog probes to determine their baseline measurement limitations and their potential for improvement using error correction.

1) **Measurement of Probe Accuracy**. We measured the errors of pendular and parallel-acting type probes using **MiniMeas**, our specially constructed probe testing apparatus. With **MiniMeas,** a probe's output could be recorded for accurately known positions of its stylus. In a typical experiment, the probe stylus was positioned in a grid pattern. The difference between the actual position of the probe's stylus and its position as reported by the probe could then be plotted. In addition to providing a baseline test of the accuracy of the probe, this plot also allowed a direct visualization of the probe's coordinate system.

2) **Identification of the Probe Coordinate System.** We continued the characterization of a probe by determining the form and parameters of its coordinate system. We found that, in the case of the two probe types tested, a rectilinear coordinate system adequately described the probe data. The optimal scales for the X and Y axes, and the angle between these axes were determined by a least squares tensor metric method. Once these parameters were known, it was a simple matter to transform the probe's coordinate system to a Cartesian one. This linear transformation completed the error correction of the probe. Transforming the probe's coordinate system to Cartesian simplifies its use with a CMM by allowing straightforward probe-CMM alignment, and arithmetic combination of CMM and probe measurements.

STATIC AND DYNAMIC PROPERTIES OF FREE STANDING BALL BARS

S. D. Phillips, B. Borchardt and T. Doiron
Precision Engineering Division
National Institute of Standards and Technology
and
J. Henry
Sheffield Measurement Corporation

This abstract is a summary of two papers which consider the static and dynamic properties of free standing ball bar systems and measurements. These systems are more widely used today than earlier methods such as magnetic ball bars which are fully described the technical literature [1]. With the release of the ANSI/ASME B89.1.12M Standard [2], *Methods for Performance Evaluation of Coordinate Measuring Machines,* in 1985, the ball bar has become an accepted method for checking the volumetric uncertainty of a CMM. In an analogous manner to other tests in the standard, the errors associated with the ball bar and stand should be less than 20% of the volumetric uncertainty of the CMM. With the development of highly accurate CMMs, with volumetric uncertainties of less than 10 μm over a work zone of a cubic meter, a more serious examination of the errors inherent in the ball bar artifact, and the measurement strategy used to measure it, is needed.

Static Properties [3]

The largest ball bar recommended by the standard is 900 mm. If we consider such a ball bar (O.D. 25 mm, I.D. 19 mm) mounted on a 900 mm rigid post (O.D. 63 mm, I.D. 38 mm) and examine the errors associated with the ball bar system, they include errors from distortions due to gravity, mechanical hysteresis, ball sphericity, vibrations, distortions due to the probe force, and thermally induced errors. It can be shown that the errors due to the sag of the ball bar from gravity are negligible. Similarly, a well clamped and mounted ball bar will have a small mechanical hysteresis. Ball sphericity errors can be reduced to approximately 0.1 μm (4 μin) by using highly spherical balls. Vibrations of the ball bar system which are excited by the forces which accelerate the carriages of the CMM to the probe approach speed can be significant. This is particularly true for touch trigger probes because the CMM must be in motion during the measurement. Analog probes which are used quasi-statically, that is they probe the part while the CMM is stationary, greatly reduce this problem since this allows time for the vibrations to dampen out. Measurements on free standing ball bar systems using touch trigger probes show that typical errors associated with vibrations are approximately 0.4 μm.

Analog probes exhibit an error not normally seen with touch trigger probes (except at very slow speeds), which involves the elastic bending of the support post. Since many analog probes apply a probing force for tens of milliseconds or longer, the ball bar system has time

to respond to this force which results in the displacement of the ball bar due to bending of the supporting post. This bending can easily be measured by placing an indicator against one ball while applying the probing force against the other ball along the direction of the ball bar axis, and observing the deflection. This deflection occurs at each end of the ball bar, and therefore the CMM will measure the ball bar length to be short by twice this value.

The largest errors which may occur during a volumetric performance test using a steel ball bar are of thermal origin. The thermal time constant for a large ball bar is over 20 minutes, hence if the ball bar is placed in a new environment with a 4°C temperature difference, at least two hours should be allowed for the system to stabilize at the 0.5 μm level. Similarly, the heat transferred by the operator upon assembly of the ball bar must be allowed time to dissipate. Once a steel ball bar has come to equilibrium it is critical to measure the temperature of the ball bar and correct for its thermal expansion during each ball bar measurement. The residual thermal error after such a correction is due to the uncertainties in the thermal expansion coefficient of the bar and of the temperature measurement. For high accuracy tests, the expansion coefficient should be known to better than 0.5 ppm/°C, and the temperature measurement to 0.05°C. Similarly the temperature range during the entire volumetric ball bar test should be less than 0.5°C.

Table 1 lists typical values for the errors discussed in this paper associated with a free standing ball bar system. Errors external to the ball bar are not considered, for example the dynamics of the CMM, or the repeatability and accuracy of the probe. For a total error budget we sum the appropriate columns of Table 1 in quadrature (root of the sum of the square of the individual terms).

TABLE 1

Typical Errors for a Free Standing Steel Ball Bar

Sizes in millimeters (inches); Errors are in micrometers (microinches)

Ball Bar Size	Supporting Post Size	System Hysteresis Error	Ball Sphericity Error	Post Bending Error	Vibration Error	Thermal Error (Steel)	Total Touch Probe	Total Analog Probe
900 long O.D. 25 I.D 19	900 high O.D. 63 I.D. 38	0.08 (3.0)	0.10 (3.9)	0.52 (20.5)	0.35 (13.8)	0.72 (28.8)	0.82 (32.3)	0.91 (35.7)
Notes				1	2	3		

Notes:

1. Post bending error applicable only to analog probes and touch trigger probes used at slow (less than 2 mm/s) speeds.
2. Vibration error is not applicable to analog probes used quasi-statically.
3. Based on a temperature change during the ball bar test of 0.5°C with uncertainty 0.05°C, and a coefficient of expansion of 11 ± 0.5 ppm/°C.

Dynamic Properties [4]

The manner in which a ball bar is measured can significantly affect its reported length. Factors which affect this measurement include machine (CMM) dynamics, probe approach speed, probe dynamics, ball bar dynamics, probing force level, and the number and distribution of probing points. CMMs which use analog probes quasi-statically, *i.e.* the CMM is brought to rest before the measurement is recorded, will avoid most of the dynamic problems. For these systems (and for touch trigger probes used at very slow speeds) it is recommended that the ball bar be measured in a two point manner, in which one point is taken on each ball at the point where the ball surface intersects the line connecting the ball centers. Under this measurement strategy the primary error is due to the probing force deflecting the supporting post of the ball bar, as described in the previous section.

We now consider touch trigger probe systems in which the measurements are taken while the CMM is in motion. CMM dynamics can usually be minimized by increasing the probe approach time during which the CMM is at constant velocity. This allows time for the machine vibrations, which are excited when the CMM carriages are accelerated to the probe approach speed, to dampen out. The measurements reported in this section have a substantial probe approach time to eliminate machine dynamic effects.

A significant effect on the measured ball bar length involves transverse bending of the ball bar. By transverse bending we mean displacement of the bar perpendicular to the bar axis due to the probing force. At first glance, transverse bending might be considered a cosine error since it is orthogonal to the axis of the ball bar; however, this is not the case. Consider a sphere defined by four points, three located on the "equator", *i.e.* perpendicular to the bar axis, and one located at the "pole", directly along the bar axis. At typical probing speeds (greater than 8 mm/s) the CMM traverses the probe pre-travel distance in a sufficiently short time that the probe triggers before the ball bar (because of its inertial mass) can significantly deflect. As the probing speed decreases the ball bar begins to bend transversely but not longitudinally (along the bar axis) because the longitudinal axis is in compression (and the supporting post is stiff). This results in the three equatorial points taken on the sphere to report a smaller radial distance from the ball axis than in the undeflected case. The working of the least squares fitting algorithm is such that the center location of the sphere is now forced out along the bar axis by an amount equal to the transverse deflection. Hence, what appeared to be a cosine error is converted, via the fitting algorithm, into a direct error. As the probe approach speed is further reduced a greater transverse deflection occurs until the restoring force equals the probing force. Thus, the ball bar will appear to grow in length over the probe approach speeds which involve transverse bending. The rate of apparent ball bar growth is further modulated by a subtle effect of touch trigger probe dynamics. The transverse bending effect is clearly seen in Figure 1 as the sharp rise in ball bar length between 4 and 2.5 mm/s. At very slow

speeds (below 2 mm/s) the probing force is applied for a sufficiently long time that longitudinal bending (bending of the supporting post) occurs, and begins to reduce the apparent length of the ball bar.

In order to avoid these dynamic effects in ball bar measurements care should be taken to insure that all probing speeds are well within the dynamically stable region, above 7 mm/s in Figure 1. For analog probes used quasi-statically the two point measurement strategy previously described is recommended to avoid transverse bending effects.

Acknowledgments: This research was funded by NIST's Automated Manufacturing Research Facility through the U.S. Navy's Manufacturing Technology Program and by the U.S. Air Force's Measurement Standards Lab.

References

1. J.B. Bryan, *Precision Engineering,* 4, p 61-69 and p 125-138 (1982).

2. ANSI/ASME B89.1.12M 1985 and 1991, *Methods for the Performance Evaluation of Coordinate Measuring Machines*, published by the American Society of Mechanical Engineers, United Engineering Center, 345 East 47th Street New York, NY 10017.

3. S. D. Phillips, B. Borchardt, T. Doiron, and J. Henry. *Properties of Free Standing Ball Bar Systems*, to be published.

4. S. D. Phillips, B. Borchardt, and J. Henry, *Dynamic Effects During Ball Bar Measurements,* to be published.

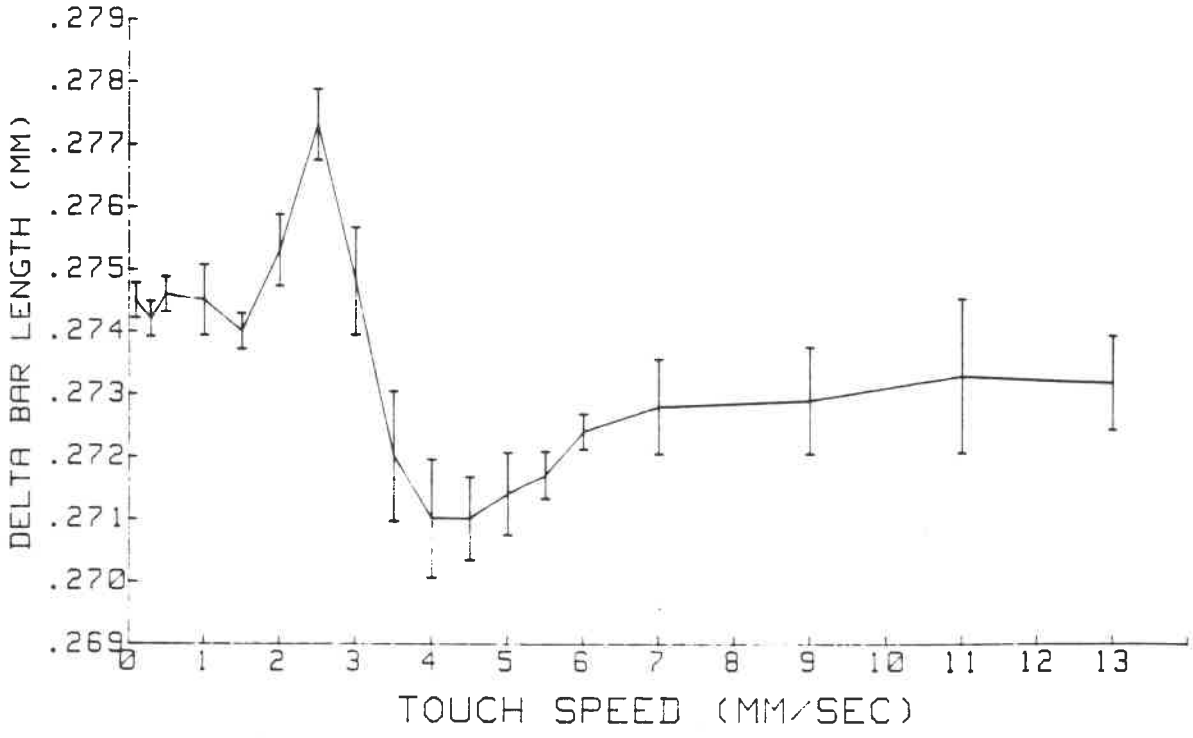


Figure 1. Ball bar length versus probe approach speed.

The Efficiency of Artifact Based Procedures to Calibrate Coordinate Measuring Machines

J. A. Soons, P. H. Schellekens
Eindhoven University of Technology
The Netherlands

1 Introduction

A variety of methods has been suggested and applied to estimate a Coordinate Measuring Machine's volumetric accuracy using artifacts. They comprise the measurement of suitable reference artifacts, such as ring gauges, bore plates, space frames and (laser interferometric) step gauges, in various locations. A fundamental advantage of these methods is that the CMM's accuracy is estimated from observed deviations in the *intended* operation of the machine. Therefore, these methods can be directly applied to software error compensated machines.

A reliable statement of a CMM's accuracy, requires the respective calibration to assess the main parametric errors which determine this accuracy (i.e. the angular, squareness, straightness and scale errors). Using linear regression techniques, a uniform approach has been developed to estimate these errors from the artifact's measurements. The same regression techniques are used regardless the nature and location(s) of the artifact(s) or the nature of the machine being studied. This includes machines with rotary tables or machines with significant finite stiffness related errors. The often complicated separation of the simultaneous parametric errors' effect on the artifact's measurements, is automatically realized by the regression technique. Therefore, artifact locations where many errors affect the measurement's result can be meaningfully used in the estimation.

From the methodology, unambiguous mathematically defined criteria are derived to evaluate the potential and efficiency of the calibration with respect to the location(s) and choice of the artifact(s) used. In addition to the evaluation of existing calibration setups, it will be shown that the derived criteria can be used in their computer aided optimization.

2 Theory

In the first stage of the proposed estimation method, a model is built of the machine's error structure, including a suitable description of its parametric errors. According to the nature of the measured feature(s), an algebraic linear relationship is established between the respective deviations and the unknown parameters which describe the machine's errors. Finally, these parameters are estimated by standard linear regression techniques[1, 2], such that the residual sum of squares between the observed and modeled errors in the measured feature(s) is minimized. The regression technique directly yields the required mathematically defined criteria to evaluate and optimize the calibration's efficiency.

2.1 Modeling a CMM's errors

The currently used models to describe a CMM's accuracy, (e.g. [3, 1, 2]) express the errors $\Delta\bar{P}$ in the measured position P of a point as a linear combination of the machine's parametric errors. For a CMM having n axes this can be formally expressed as:

$$\Delta\bar{P} = \sum_{i=1}^{n} ({}_{tl}F_i\,\bar{E}_i) + \bar{E}_{tl} \qquad (1)$$

In equation 1 the vector $\bar{E}_i$ contains the three rotation and three translation errors introduced in axis i and its surrounding components. The 3×6 matrix ${}_{tl}F_i$ describes how these parametric errors affect the errors $\Delta\bar{P}$ in the measured position of a point. This matrix is completely defined by the nominal geometry of the machine, the used probe and the nominal position P of the measured point. Vector $\bar{E}_{tl}$ contains the translation errors introduced by the probe system.

In order to obtain a model which describes the CMM's accuracy at a certain time and place, it is necessary to relate the parametric errors $E_{i,j}$ to the status of the machine (e.g. the position of its axes). In many studies, essentially a linear interpolation is used between the measured or estimated values of these errors at discrete positions of the respective machine's axis. In our approach, this description is generalized by modeling each parametric error $E_{i,j}$ as an estimated linear combination of known functions. These functions are defined on the position of the machine's axes and other relevant variables (e.g. the output of certain temperature sensors).

$$E_{i,j}(\bar{x}) = \sum_{k=1}^{m} p_k(\bar{x})\beta_k \tag{2}$$

In equation 2 $p_k(\bar{x})$ is a known function defined on the relevant independent variables contained in vector $\bar{x}$. Good results have been obtained with piecewise polynomials[2]. Due to finite stiffness effects, a parametric error can be related to the position of axes other than that of the respective kinematic element[2].

2.2 Estimation of a CMM's parametric errors using artifacts

In this section we consider the estimation of the parameters $\bar{\beta}$ which describe the machine's parametric errors, using deviations observed when measuring artifacts. First we need a suitable model which relates these deviations to the errors in the realized (relative) position of the machine's probe tip when measuring these artifacts. Using the following approximation, a linear relation is obtained which greatly simplifies the subsequent estimation process.

> The difference between the measured and actual location of an artifact *due to the CMM's errors* has a negligible effect on the measurement errors of its features.

Usually this approximation only yields problems for CMM's whose errors change significantly over intervals of the same order as its errors. Also if the calibration has to predict the CMM's errors with an extremely high relative accuracy, a nonlinear approach may be necessary.

As an example, we consider an artifact whose length is calculated as the distance between the measured points P_1 and P_2. Using the above-mentioned approximation, the observed error ΔL in its measured length L_{meas} can be modeled as a *known* linear combination of the measurement errors $\Delta\vec{P}_1$ and $\Delta\vec{P}_2$:

$$\Delta L \approx \frac{(\vec{P}_2 - \vec{P}_1)}{L_{meas}} \cdot (\Delta\vec{P}_2 - \Delta\vec{P}_1) \tag{3}$$

Although this result (i.e. the neglect of cosine errors) may seem trivial, it is essential that a linear relation is obtained. Also the machine's error model (see equation 1) and the parametric error's model (see equation 2) are linear in respectively the parametric errors $\bar{E}_i$ and the unknown parameters $\bar{\beta}$ which describe these errors. Thus the appropriate combination of these models, yields a relation in which the measurement error ΔA_i of a feature A_i is expressed as a (known) linear combination of the parameters $\bar{\beta}$:

$$\Delta A_i = \sum_{k=1}^{p} {}_i g_k(\bar{x}_i)\beta_k + \gamma_i \tag{4}$$

Here the known function ${}_ig_k(\bar{x}_i)$ describes the effect of parameter β_k on the artifact's measurement error. In the case of calibrations which use artifacts of unknown constant dimensions, the vector $\bar{\beta}$ has to be expanded to include these dimensions.

Thus a calibration based on the measurement of m artifact features can be presented mathematically as:

$$\Delta\bar{A} = X\bar{\beta} + \bar{\gamma} \tag{5}$$

Here $\bar{\gamma}$ is the vector of random errors assumed to be independent and identically distributed with mean 0 and variance σ^2. Row i of the $m \times p$ matrix X contains the function values ${}_ig_1(\bar{x}_i), \ldots, {}_ig_p(\bar{x}_i)$. From equation 5 the unknown parameters $\bar{\beta}$ can be estimated using linear least-squares analysis[4]:

$$\hat{\bar{\beta}} = \left(X^T X\right)^{-1} X^T \Delta\bar{A} \tag{6}$$

2.3 Mathematical description of a Calibration's Efficiency

The efficiency and potential of a calibration with respect to the choice and location(s) of the artifact(s) used, can be assessed by two criteria:

- the assembly of error parameters which can be estimated from the performance evaluation's observations
- the accuracy of these estimators, given the repeatability of an artifact's measurement.

The (proposed) model for the CMM's systematic errors affects these criteria in a structural manner. Hence every calibration is a compromise between its efficiency in the estimation of modeled errors and its ability to detect errors not implemented in the proposed model. With respect to the calibration of CMMs, there is a priori information available regarding the machine's kinematic model. Further specification of the proposed model (i.e. the complexity of the models for the parametric errors) is a compromise between the allowed calibration's effort and the final model's capability to describe irregular parametric errors.

From equation 6, the calibration's potential to estimate the unknown parameters $\bar{\beta}$ can be assessed by the following rule[4]:

> *a linear combination $\bar{\lambda}^T\bar{\beta}$ of the parameters $\bar{\beta}$ can be assessed from the calibration's observations, if $\bar{\lambda}^T$ is part of the vectorspace spanned by the rows of X.*

In case all the parameters $\bar{\beta}$ can be estimated, the calibration's efficiency can be further assessed by[4]:

- the variance-covariance matrix of the estimated parameters $\hat{\bar{\beta}}$:

$$\mathrm{Var}(\hat{\bar{\beta}}) = \sigma^2(\mathrm{X^T X})^{-1} \tag{7}$$

- the variance of the predicted deviation when measuring an artifact at a certain location:

$$\mathrm{Var}(\Delta\mathrm{A} \mid \bar{\mathrm{x}} = \bar{\mathrm{x}}_\mathrm{i}) = \sigma^2{}_\mathrm{i}\bar{\mathrm{g}}(\bar{\mathrm{x}}_\mathrm{i})^\mathrm{T}(\mathrm{X^T X})^{-1}{}_\mathrm{i}\bar{\mathrm{g}}(\bar{\mathrm{x}}_\mathrm{i}) \tag{8}$$

Both 7 and 8 depend on the design of the calibration setup only through the $p \times p$ matrix $(X^T X)^{-1}$. This suggests that a good calibration setup will be one that makes this matrix small. Since there is no unique size ordering of $p \times p$ matrices, various real-valued functions have been suggested as measures of 'smallness'. Some criteria[5] focus on the quality of the estimated parameters

(minimization of the determinant, trace or maximum eigenvalue of $(X^TX)^{-1}$). Others address the quality of the estimated response surface for the errors in the artifact's measurements (minimization maximum or average variance predicted errors). Although the comparison of existing calibration setups by these criteria is quite straightforward, optimization is a very complicated process. Major computational difficulties are encountered due to the complex nature of the object function and the large number of design variables which describe a calibration setup.

3 Applications

In this section the efficiency criteria are applied to calibrate a bridge-type CMM using a relatively small number of length measurements. This setup presents an extreme of the efficiency compromise explained in section 2.3. Since the choice and position of the measured lengths is optimized according to the proposed error model, the efficiency in the estimation of the unknown parameters $\bar{\beta}$ is high. Thus only a small number of measurements is required. However a too rigorous optimization, will significantly reduce the possibilities to detect CMM errors not implemented in the model. Yet in most cases, it is possible to predict the nature of the errors which have to be determined during the calibration. For example if, due to finite stiffness effects, a certain parametric error is dependent on the position of one or more axes.

In order to facilitate the evaluation of the method's potential, the same length reference (a laser interferometer) was used during both the calibration and the subsequent tests of its validity. Not included in the calibration were the errors introduced by the CMM's probe system. Angular errors which affect the measurement's accuracy only through the used probe offset (in our case Z-axis roll and X-axis yaw) where also not estimated. As with probing errors, dedicated tests, using different probe offsets, are better suited to assess these errors.

Disregarding the used probe offset as a design parameter, a calibration using n length measurements requires the simultaneous optimization of $n \times 6$ variables (5 variables for the position and orientation of each length in the machine's workspace and 1 variable for the choice of this length). In the chosen approach[5] to this rather difficult problem, a discrete design space X is defined of r so-called candidates $x_1, \ldots, x_r$. Each candidate represents a length measured at a certain location in the CMM's workspace. The adapted (exchange) algorithm starts with an initial random design of n candidates (not necessary distinct). This is followed by a number of excursions whereby the design is improved until no further progress can be made. Each excursion is a series of additions or subtractions of one or more candidates to the current collection of points, eventually returning to a possibly new and better n-point design. A design was considered to be more efficient, if the determinant of the parameters' variance-covariance matrix (see relation 7) was smaller. Essentially this so-called D-criterion tries to minimize the content of the (multi-dimensional) confidence region associated with the estimated parameters $\bar{\beta}$.

The number of lengths used in this calibration was chosen as 50. In order to determine the trend of the various errors, low order (quadratic or linear) piecewise polynomials were used in their proposed model. This yields a total of 32 unknown parameters to be estimated. The optimization algorithm was allowed to choose from 'artifacts' of 350, 500 and 800 mm length.

In order to evaluate the calibration's accuracy, we use a laser interferometer to measure various lengths in combination with a mirror mounted on a motorized rotary table (see Figure 1). A large retroreflector is attached to the machine's ram and the interferometer is placed between the laser source and the mirror. The mirror setup is (software) aligned by measuring the location of the reflected beam for several angles of the rotary table. Each location is estimated from the machine's scale readings at two different, manually tuned, positions where the laser beam is optimally returned

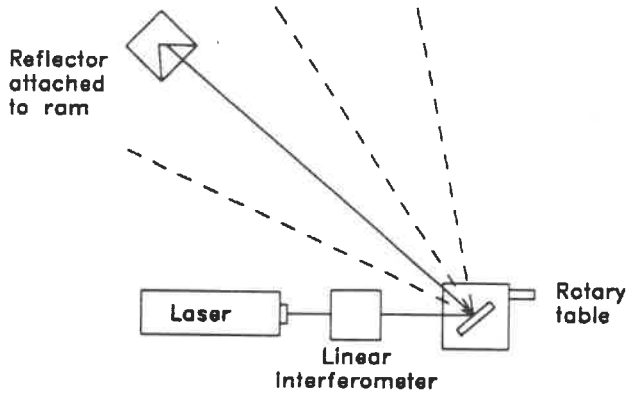


Figure 1: Experimental setup to measure 'many' lengths

to the laser head. From this information the host computer is able to calculate the position and direction of the reflected laser beam for any angle of the rotary table. In a certain fixed position of the rotary table the machine's ram is now automatically moved along the reflected laser beam. At certain positions the machine is stopped and its indicated travel is compared with the respective laser reading. After a complete back and forth movement, the mirror is automatically rotated by a certain angle and the process is repeated. Thus a large number of lengths can be measured fully automatically.

Note that for this application the accuracy requirements of the rotary table are rather low. Since the position of the mirror is fixed during each length's measurement, errors in its location only affect the measurement's accuracy by cosine errors. Similarly the accuracy requirements of the mirror surface and the retroreflector are rather low. A drawback of the setup is the increased distance between the interferometer and the starting point of the length measurements, due to mirror and rotary table. Significant 'deadpath' errors might occur in case of an unstable environment and slow execution of the back and forth movement.

Using this setup a total of roughly 2000 lengths were measured throughout the complete CMM's workspace. Figure 2A presents the observed errors in the measurement of these lengths. In Figures 2B, 2C, and 2D the residuals between the predicted and observed length deviations of this dataset are depicted for models obtained with respectively the small number of length measurements in optimized locations, conventional measurement equipment (a laser interferometer and electronic levels) and the measurement of a separate set of also roughly 2000 length measurements. The latter so-called 'brute-force' calibration indicates the accuracy which can be obtained by artifact based calibration procedures.

4 Conclusion

A general methodology has been presented to estimate the quasi-static errors of CMMs using artifact measurements. The use of standard regression techniques in the estimation of the CMM's parametric errors, yields an uniform approach with respect to the nature of the machine being studied and the choice and location(s) of the artifact(s) measured. Using standard statistical experimental design techniques, the potential and efficiency of artifact based calibrations can be unambiguously assessed and optimized. The methodology has been applied to calibrate a three axis CMM using a limited number of length measurements in optimized locations. The predictive power of the estimated CMM's error model is comparable with that obtained using conventional calibration techniques.

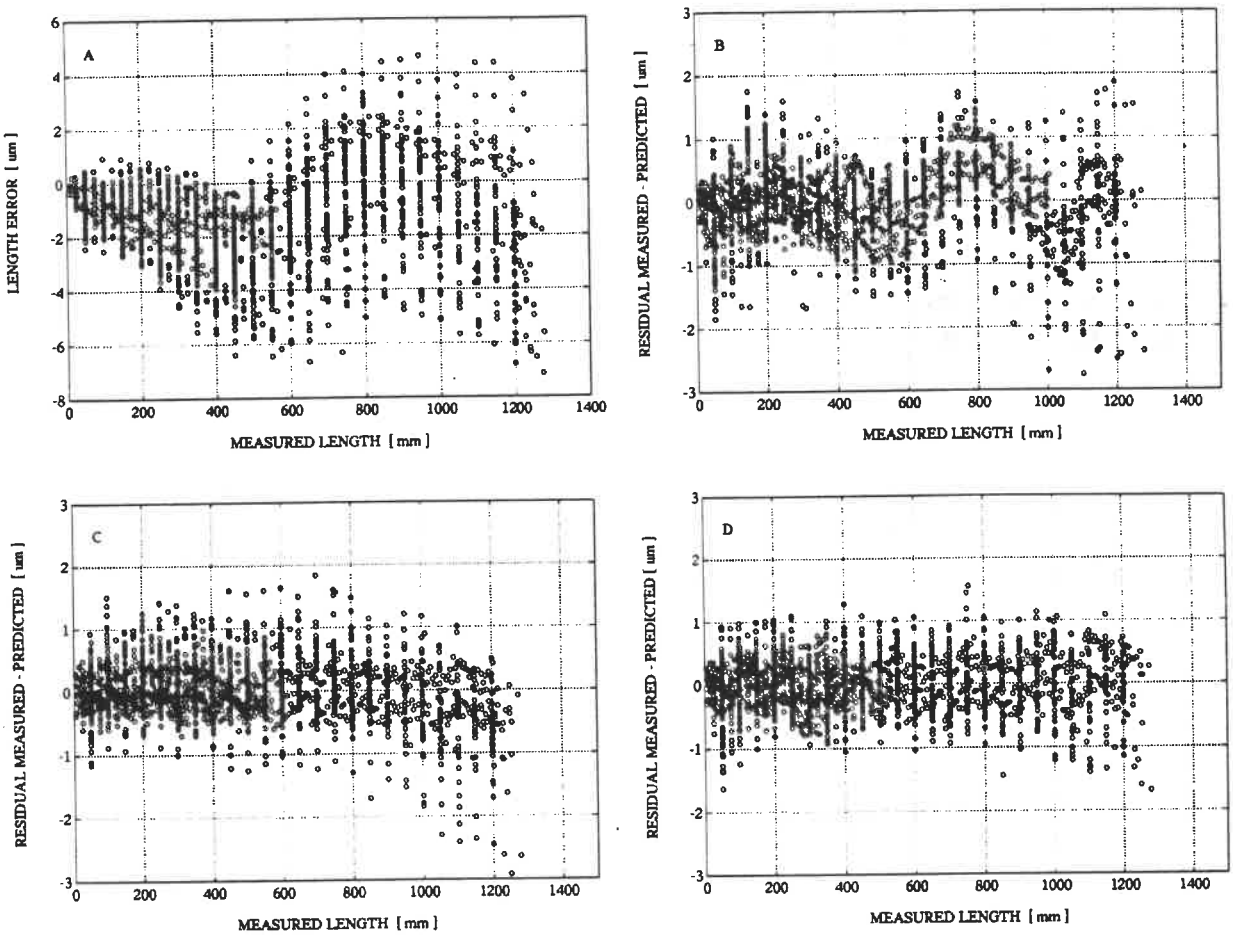


Figure 2: Errors in the observed lengths measured by the CMM, and the predictive capability of the various calibration methods

References

[1] Ferreira P. M. and Liu C. R. A contribution to the analysis and compensation of the geometric error of a machine center. In *CIRP Annals*, number 35, pages 259–262, 1986.

[2] Soons J. A., Theuws F. C., and Schellekens P. H. Modelling the errors of multi axis machines: A general methodology. *Precision Engineering*, 13, 1991.

[3] Hocken R. et al. Three dimensional metrology. In *CIRP Annals*, number 26, pages 403–408, 1977.

[4] Kshirsagar A. M. *A Course in Linear Models.* Marcel Dekker Inc, New York, 1981.

[5] St. John R. C. and Draper N. R. D-optimality for regresion designs: A review. *Technometrics*, 17(1):15–23, 1975.

Error Analysis for X-Y Interferometric Metrology

Norman Bobroff
IBM Thomas J. Watson Research Ctr.
PO Box 218, Yorktown Hts, NY 10598

A model of two dimensional in-plane metrology using plane mirror interferometry is developed based on the archetype measuring configuration of Fig. 1. The purpose of the model is to define and study the key factors limiting the accuracy of this system. Behavior of the metrology system is treated in two steps. Geometric errors are studied under the assumption that component motions and transducers are linear, but not necessarily in alignment. Abbe offsets depend on higher order terms and are neglected in the first order analysis. This study provides some interesting clarifications of which alignment errors are most important and what is meant by cosine error and alignment error. In the second step, a detailed numerical model of the system including periodic, random and systematic errors is developed. Accurate data on the performance of the interferometric transducers is essential to an accurate metrology model which can be used to assess self-calibration techniques. A significant component of the second step is to characterize interferometer accuracy, and some new methods for doing this are presented.

Figures 2a, b, and c show the metrology vectors associated with the x-y system of Fig. 1. Fig. 2a. is a plan view of both mirrors and the x and y measurement vectors. Tx is the direction of travel of the x stage coordinate, and is the line of intersection of the plane of the x interferometer mirror with the base travel plane of the stage. Ix is the interferometer beam direction used to measure the amount of travel along Tx, and Nx is the normal to the measurement mirror. The vectors are unit vectors so that only two components are independent. An analogous set of definitions applies to the y axis, and the plane of travel contains Tx and Ty although these vectors may not be orthogonal. The three dimensional nature of the these vectors is emphasized in Fig. 2b, which defines the two independent direction cosines β and γ as projections of the vectors onto the y-z plane. The direction cosine projection is used in Fig. 2c to show the three x axis vectors on a two dimensional plot. The distance measured by the interferometer when the stage is moved one unit along T is $d = (N \cdot T)(N \cdot I)$. The first inner product is the cosine error, while the second is the interferometer alignment error. Note that there is no term in $I \cdot T$. This relation shows that the transducer axis of an interferometer in the plane mirror configuration is the mirror normal not the

interferometer beam, I. For comparison, a retroreflector target causes the transducer axis to be the interferometer beam.

The six vectors of Figs. 2 lead to a view of metrology alignment based on the vector pairs (N_x, N_y), (I_x, I_y), and (T_x, T_y) each of which defines a parallelogram. Alignment of the metrology stage has the goal of making the parallelograms coplanar, and minimizing deviations from right angles. Calibration is the effort to measure the orientation in space of the parallelograms, as well as their internal angles. There are 18 parameters associated with the alignment and calibration of the six vectors, however they are not all independent. The travel vectors define the x-y plane and T_x is taken to make the same angle Φ with the x axis that T_y makes with the y axis. Normalization of the vectors, orthogonality $d = (N_x \bullet T_y) = (N_y \bullet T_x) = 0$, , and symmetry constrain the system to seven independent vector alignments. Using the direction cosine perspective, these are the α or x axis component of I_y, the β or y axis component of T_x and I_x, and the γ or z components of N_x, N_y, I_x, and I_y.

It is interesting to see how these degrees of freedom effect system accuracy. Consider a vector of magnitude r and angle α in the plane of an orthogonal x-y coordinate system which is transformed to a nonorthogonal system of the measurement tool with angle 2Φ between axes T_x and T_y. The Cartesian vector components transform as

$$(r\cos\alpha,\ r\sin\alpha) \rightarrow \frac{1}{(T_x X T_y)}\ (r\cos(\alpha-\Phi),\ r\sin(\alpha+\Phi))$$

and the magnitude measured by a metrology tool with no scale error, $(N \bullet T)(N \bullet I) = 0$, is

$$r'^2 = \frac{r^2}{(T_x X T_y)^2}\ [1 + (T_x \bullet T_y)\ \sin 2\alpha]\ ,$$

while the angle α' is related to α by

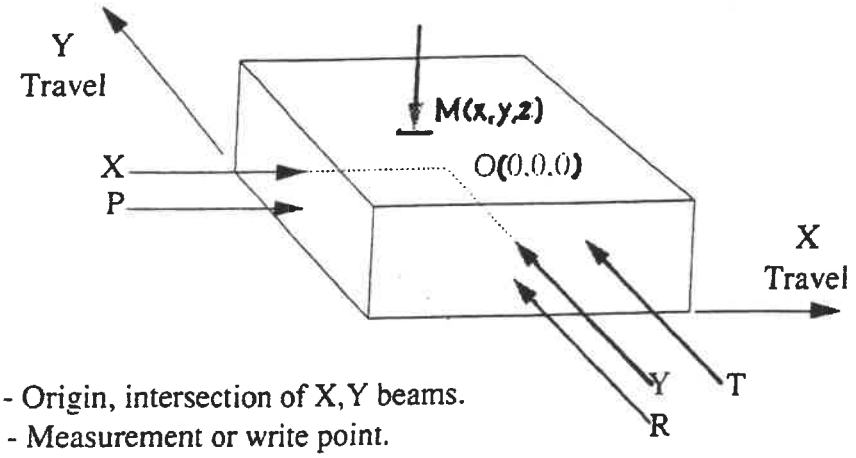
$$\tan(\alpha') = \frac{[\tan(\Phi) + \tan(\alpha)]}{[1 + \tan(\Phi)\tan(\alpha)]}$$

In practice, cosine and alignment errors combine with the coordinate transformation so that each axis has a net scale factor of the form $(N \bullet T)(N \bullet I)/(T_x X T_y)$. The vector constraint relations and a small Φ approximation simplify the net axis scale to

$$(N \bullet I)[1 - 1/2((T_x X T_y) \bullet N_x)^2]\ ,$$

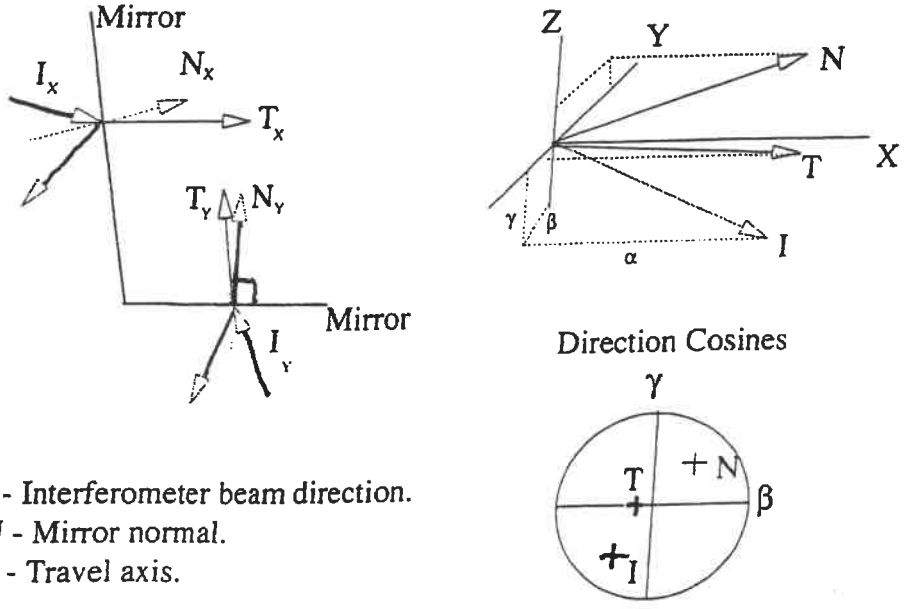
and a similar relation holds for the y axis. The explicit cosine error component ($N_x \bullet T_x$) does not appear, only the out of plane component of N affects the measurement of distances and angles. This component of N is often the most difficult to control and characterize in the mounting of stick mirrors for these stages. Nonorthogonality is the only first order term which contributes to errors in distance and angle.

A model of stage errors is developed based on geometry and interferometer errors. The model incorporates polynomial descriptions of the stage pitch and roll, and interferometer mirror figure. Plane mirror interferometers are assumed to monitor the x and y axes, and optionally the pitch, roll, and yaw axes. Abbe offsets take into account the fact that the interferometer axes may not have a virtual intersection in space. A large number of sources of systematic and random errors are incorporated in the model. One interesting interferometer transducer error included in the model is the influence of wavefront aberrations. The interferometer beams are not uniphase, and generally have of the order of $\lambda/8$ residual curvature and spherical aberration. In the presence of alignment errors, or stages which rotate the interferometer mirrors, the interfering regions on the interferometer wavefronts can depend on stage position. Another major portion of the work is characterization of errors introduced by the receivers and phase meters. Intercomparison of receivers and phase meters which view the same optical input can show periodic phase measurment errors, as well as drifts due to variable delay times in the electronics. Fixed phase differences, introduced optically between phase meters provide additional information on phase meter accuracy. Intercomparison studies are described in more detail in the paper by Greier et al in this proceedings. One conclusion of this work is that the random contribution to errors is about 5nm at the 1σ confidence level for a system operating in vacuum and making use of commercial interferometer systems. Furthermore, the component contributions to random error from the phase meters, digital resolution, electronic noise, and Abbe errors are of the similar magnitude order so that substantial error reduction is difficult.



O - Origin, intersection of X,Y beams.
M - Measurement or write point.
P - Pitch measurement.
R - Roll measurement.
T - Yaw measurement.

Fig. 1.Stage with five axis interferometric metrology.



I - Interferometer beam direction.
N - Mirror normal.
T - Travel axis.

Fig. 2. a) Each metrology axis has three associated vectors; I - interferometer beam, T - Travel, N - mirror normal. b) The three dimensional orientation of vectors I, T, N. c) Direction cosine view of I, T, N.

Electronic Limitations in Phase Meters for Heterodyne Interferometry

Nile Oldham, Paul Hetrick, John Kramar
William Penzes, Tom Wheatley, and Clayton Teague
National Institute of Standards and Technology
Gaithersburg, MD 20899

Reasonable attention has been given to the fidelity of the process by which heterodyne interferometers convert optical path difference between beams that have traversed a test leg and a reference leg, respectively, to a phase difference between electrical signals from the test and reference photodetectors. This paper reports on a study of the next step; to obtain a quantitative result from these signals by measuring the electrical phase difference between the two photodetector signals.

In the least demanding applications, where the test reflector is moved and then stopped (with respect to the reference reflector), a fixed phase relationship exists between the two electrical signals at the time of measurement. Commercial phase angle meters operate at signal frequencies out to about 500 kHz, with angular uncertainties of ±0.2° and measurement speeds on the order of one reading per second. The recent advent of commercial time interval analyzers has made it possible to make phase angle measurements at the higher frequencies and update rates required by heterodyne interferometry.

The accuracy of the phase angle measurement depends on the time resolution, linearity, and stability of the time interval analyzer. In the case of a 1-MHz signal frequency, a precision of 1-ns is required to resolve 1/1000 of a period (0.36°). Fig. 1 shows the phase resolution that can be achieved at signal frequencies from 2 kHz to 20 MHz with time resolutions from 10 ps to 100 ns. There are a number of commercial time interval analyzers with the ability to resolve 1 ns or less. Their maximum update rates range from 1 kHz to 5 MHz with "single shot" time resolutions from 10 ps to 1 ns (resolution can generally be increased with averaging).

Strictly speaking, the phase angle between two signals is defined only when those signals are the same frequency; however, it is useful in interferometer applications to measure the apparent phase angle between two signals of different frequency. If the test reflector is moved to produce a signal that is changing in phase relative to the reference signal, the test signal may be expressed as:

$$V_t = A \cos (wt + \phi(t)),$$

where w = frequency of the reference and $\phi(t)$ is the rate of change of phase.

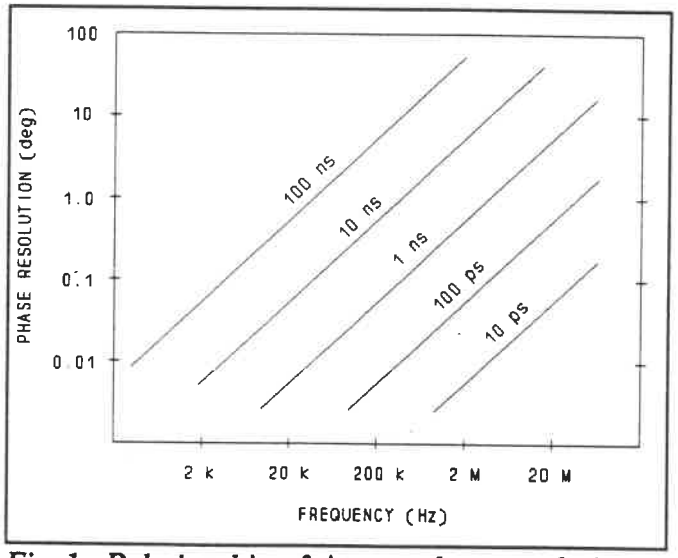


Fig. 1. Relationship of time to phase resolution at frequencies from 2 kHz to 20 MHz.

It can be shown that $\phi(t)$ is equivalent to a frequency term that, depending on the direction of phase shift, either increases or decreases the frequency of the test signal [1]. For example, if the reference signal is 1 MHz and the test signal is phase shifted at a rate of $180°/\mu s$, then the frequency of the test signal will be 1.5 MHz.

Two function generators with a common reference oscillator can be configured to produce relatively pure sine waves that are adjustable in phase angle. In one type of generator, the phase relationship between the output signal and the reference oscillator is adjustable in 0.1° steps. Thus, if two of these generators are connected to a common reference oscillator, it is possible to program the phase shift between the two output signals by adjusting the phase relationship of one or both signals referenced to the oscillator. Tests of the timing linearity of signals produced by this technique indicated that, for this particular function generator, the "time delay" nonlinearity between the reference oscillator and the output signal is less than 10 ps [2]. This corresponds to a phase uncertainty of less than 0.01° at 2 MHz. Based on this performance, a dual channel source (consisting of two function generators) was assembled to test a number of commercial phase meters and time interval analyzers from 250 kHz to 20 MHz.

Static tests (both signals at the same frequency) were performed on a number of different analyzers configured as phase meters. The results of tests performed at 2 MHz on two of these instruments are shown in Fig. 2. Based on tests at other frequencies, an experimental figure for the phase uncertainty (due to nonlinearity and random scatter) of a typical time interval analyzer used to measure phase is given by: $\pm(0.05 + 0.1F)°$, where F is the test frequency in MHz.

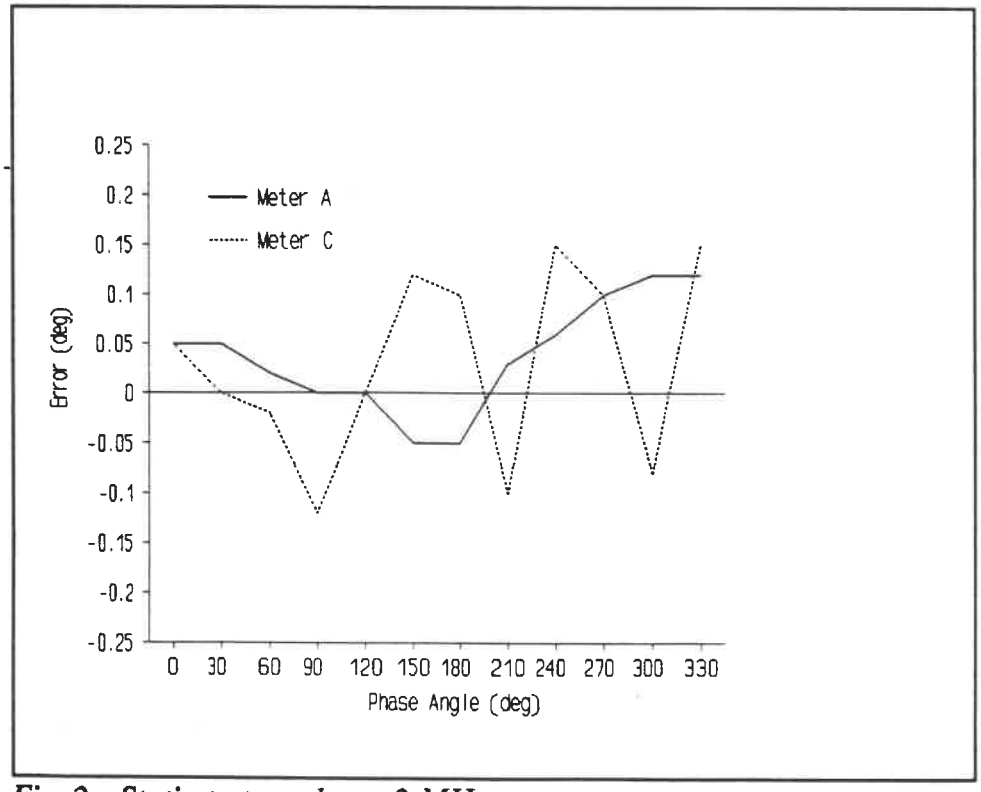


Fig. 2. Static test results at 2 MHz.

Dynamic tests were performed on two of the time interval analyzers that are capable of update rates in excess of 2 MHz. In these tests the reference signal was 2 MHz and the test signal (which simulates the Doppler shifted signal from the moving test reflector) was about 2.2 MHz giving a linear phase angle rate-of-change of 30°/μs. The dynamic nonlinearities shown in Fig. 3 are based on time interval measurements made on-the-fly at 500 ns intervals.

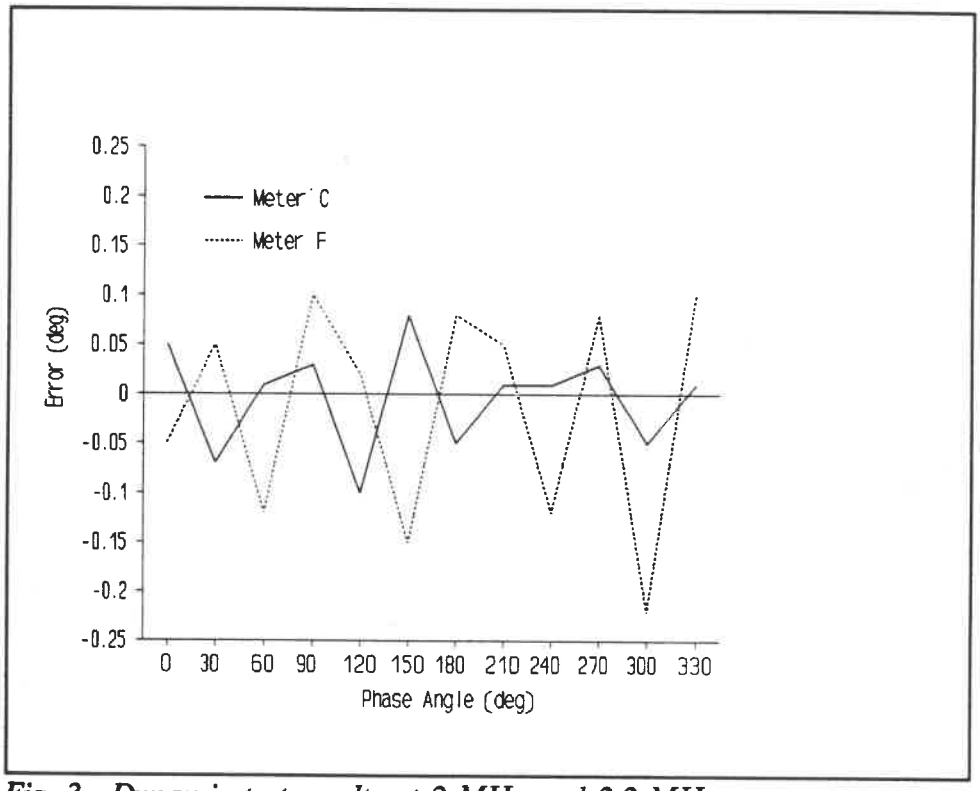


Fig. 3. Dynamic test results at 2 MHz and 2.2 MHz.

The dynamic performance of several of the newer time interval analyzers is of particular interest to users of heterodyne interferometers. In this application electrical signals, proportional to mechanical displacement can be processed to measure the position of a moving object in microsecond time regimes. The measurements reported in this paper show that, electronic limitations imposed by the phase measurement correspond to timing errors of about 350 ps (±0.25° at 2 MHz) which will limit the precision of a simple Michelson interferometer, to about 1/3000th of a wavelength at a reference signal frequency of 2 MHz.

REFERENCES:

1. Teague, E.C., "Nanometrology," Conf. Proc. of Scanned Probe Microscopy; STM and Beyond, Santa Barbara, CA, Jan. 1991, published by APS.

2. Oldham, N.M. and Hetrick, P.S., "High-Frequency, High-Speed Phase Angle Measurements and Standards," Conf. Proc. of the 1991 NCSL Workshop and Symposium, Albuquerque, NM, August 19-23, 1991, pp 252-256.

AN INTERFEROMETRIC INSTRUMENT FOR THE METROLOGY OF MICROMECHANISMS

Lowell P. Howard, Robert J. Hocken
Precision Engineering Laboratory
University of North Carolina at Charlotte
Charlotte, NC 28223

Introduction

The development of Micromechanisms has created the requirement for metrological instruments to measure their performance. Examples of practical mechanisms requiring metrology are the small cantilevered structures used in scanning probe microscopes. Two common schemes for measuring the displacement of these probe tips are the optical lever[1,2] and the fiber optic interferometer.[3,4,5] Both methods require a skilled operator for alignment, average over an ill-defined area of the surface and make it difficult to pinpoint the exact measurement area on the micromechanism. The fiber optic interferometer also physically obstructs access to the image of the surface. This paper describes an instrument designed to allow us to image a surface and make wide bandwidth vertical displacement measurements at a location and over an area of our choice. Our system allows the operator to image the cantilever (with the measuring point indicated by a fiducial mark) and then measure displacements with bandwidth exceeding 50 kHz. In a first experiment we measured step responses of electrostatically actuated SiO_2 microcantilevers.

Instrument Design

Our instrument utilizes a Zeiss microscopic interferometer as a test bed for our sensing scheme. This microscope consists of an objective lens/reference mirror turret, a sample translation/tilt stage and controls for rotation and spacing of the interference fringes. Our modifications involved the light source and detection scheme (Fig. 1).

Light Source: The original system used a green Thallium lamp with an emission wavelength of 0.54 μm. This type of extended source is ideal for visual observation and photography because it's low spatial coherence greatly reduces speckle noise. Also, it's green emission wavelength would perfectly match the spectral response of most photomultiplier tubes. This source was not used however, because of its 60 Hz AC operation resulting in 120 Hz intensity variations. We decided not to modify the tube for DC operation because of an internal starter circuit and the fact that replacement tubes were becoming extremely scarce.

Our present source is a HeNe laser that has been spatially filtered and collimated. The beam is delivered to the interferometer through an aperture made available to us by removing the condensing optics used for the Thallium source. Speckle noise from the interferometer optics is evident but does not influence the response to AC (cantilever) signals as the speckle pattern does not change and may be regarded as DC.

Detection: The original Zeiss instrument used a 35 mm camera body to provide photographic record of fringe patterns. Our modification utilizes a fiber optic cable placed in the real image plane, instead of photographic film. As a cantilever vibrates, its interference fringe pattern sweeps across the aperture of the fiber, sending light to a photomultiplier tube. Thus, a 1 mm circular fiber, (NA=0.8) when used with a 60x objective lens, maps to a 17 μm diameter circle on the object. Care must be taken to operate the interferometer about its linear region near the quadrature point. This operating point is where the optical path difference is equivalent to a phase difference of $\pi/2$. When observing a vibrating object, the total phase deviation must be kept small (0.2 rad) or energy will be lost in higher order sidebands.[6] This is best observed directly on a spectrum analyzer. In our experiments, the quadrature point was set by translating a wedge plate inside the microscope so that the output was halfway between the fringe minima and maxima.

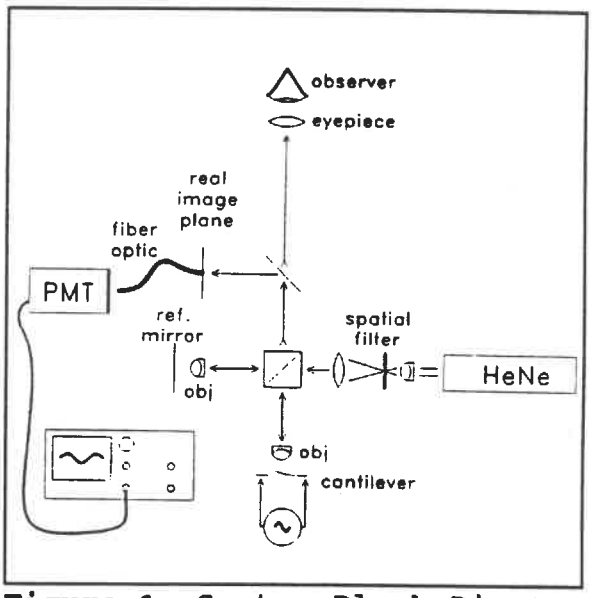


Figure 1. System Block Diagram

Cantilevers

Microcantilevers were fabricated from thermally grown SiO_2 on a <100> silicon substrate. Each cantilever is 200 μm in length, 20 μm wide and 1 μm thick. The cantilevers are grouped in arrays of five, spaced 20 μm apart, with 5 cantilever arrays per chip. Each cantilever has a chromium electrode (150 Å thick) on its tip, 4 μm wide, that makes up one-half of the electrostatic drive mechanism. The other half of the driver is a chrome metal pad spaced 2 μm from the tip of the cantilever. The cantilever is released from the substrate by an anisotropic etch that leaves the cantilevers hanging over a pyramidally shaped pit made from intersecting <111> planes. When the cantilevers are released, due to residual stresses in the thermally grown oxide and the evaporated chrome, they bend upward, leaving a vertical gap of about 3 μm between electrodes on the tip and on the edge of the etch pit. The two electrodes serve as a fringing field capacitor and applying a potential to them results in a displacement of the cantilever tip. Note that this type of arrangement does not allow the electrodes to touch. Fig. 2 shows one cantilever array with the electrode labeled.

Experimental

Noise: A RMS noise level of 8.3 nm was measured with no bandlimiting present. System bandwidth is determined by the response of the photomultiplier tube, which could in theory be well beyond several MHz. A noise spectrum is not yet available but it would most probably contain a strong 1/f laser noise component. To reduce noise while still retaining bandwidth would require modulating the laser signal at a frequency greater than 500 kHz,

possibly with an acousto-optic modulator in conjunction with a high speed lock-in amplifier.

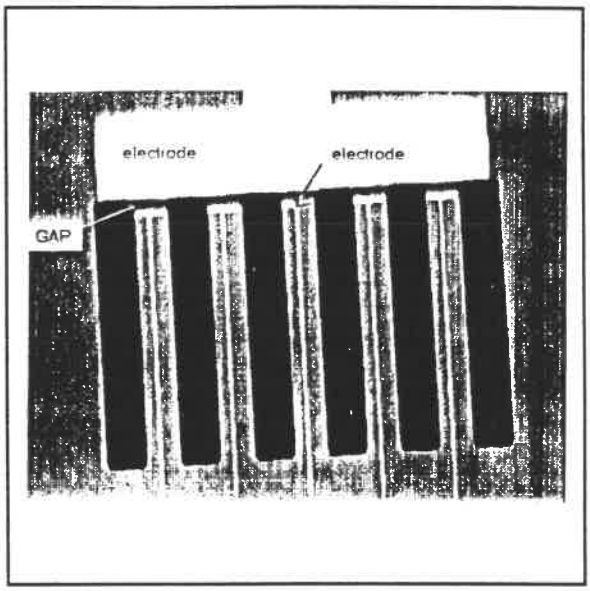


Figure 2. Microcantilever Array

Step Response: Cantilever vibrations were observed by applying a 30 V step function to the electrodes and recording the resulting oscillations on a digital oscilloscope (Fig. 3). The step function was generated using a square wave oscillator driving a MOSFET in a common source amplifier configuration. The resulting damped sinusoid showed a resonant frequency of approximately 42 kHz.

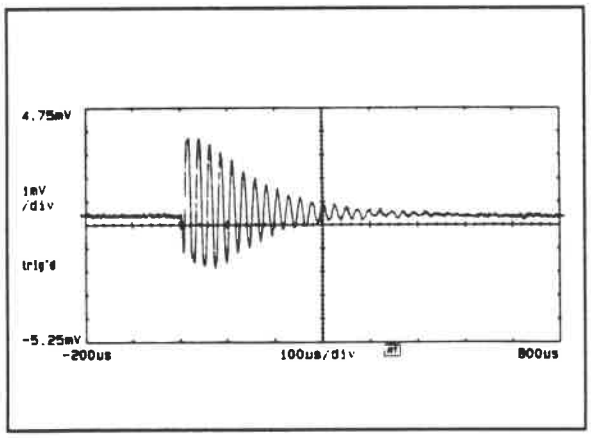


Figure 3. Step Response

Frequency Response: The internal reference oscillator of a computer controlled lock-in amplifier was amplified to a level of 30 V_{p-p} using two operational amplifiers in a bridged configuration. Under computer control, the reference oscillator was swept from 10 kHz to 50 kHz at 1 kHz intervals. At each measurement point, the lock-in was instructed to auto-phase adjust for maximum signal strength. Bode plots were generated by collecting data for reference oscillator frequency, amplitude, and phase at each point (Figs. 4 and 5).

The actual resonance frequency is at 42 kHz, twice the drive frequency due to the drive waveform swinging both positive and negative. Future work will use a DC bias voltage to prevent this, and reduce amplifier distortion at the zero crossings. The nature of the 84 kHz harmonic is unknown at this time, a spectral analysis of the drive waveform should reveal the source of the anomaly.

Spatial Resolution: A NIST calibrated AR chrome line standard was used to determine the spatial resolution of the instrument. The line standard was translated in one direction using a hand-cranked 40 to 1 reduction gear coupled to the stage micrometer screw through a flexible coupling. Stage displacement was measured using a LVDT gauge driving the X input of a digital oscilloscope. PMT output was recorded on the Y channel to provide a record of the modulation transfer function (Fig. 6).

Conclusions: Interferometry for micromechanisms by sampling a vibrating interference pattern is an effective technique for analysis of vibrating micromechanisms. Micron scale

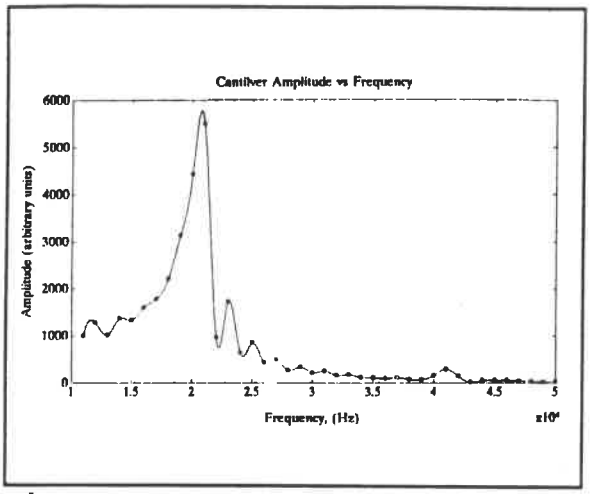


Figure 4. Bode Plot-Amplitude

devices may be quickly imaged, inspected and measured using this technique. Future work will concentrate on reducing laser noise levels through use of high frequency modulation and synchronous detection methods.

Acknowledgments: This work was funded by the National Science Foundation. The authors would like to thank C. Lofgren, C. Roberson and A. vanMeeteren for their skills and support, and R. Larrabee of NIST for the loan of the photomask standard.

References

1. R. C. Barrett and C. F. Quate, "Imaging Polished Sapphire With Atomic Force Microscopy", J. Vac. Sci. Technol. A, Vol. 8, No. 1, PP. 400-402, 1990.

2. G. Meyer and N. Amer, "Simultaneous Measurement of Lateral and Normal Forces With An Optical-Beam-Deflection Atomic Force Microscope", Appl. Phys. Lett., Vol. 57, No. 20, PP. 2089-2091, 1990.

3. P. J. Mulhern, T. Hubbard, C. S. Arnold, B. L. Blackford, and M. H. Jericho, "A Scanning Force Microscope With A Fiber-Optic Interferometer Displacement Sensor", Rev. Sci. Instrum., Vol. 62, No. 5, PP. 1280-1284, 1991.

4. S. Breen, B. E. Paton, B. L. Blackford and M. H. Jericho, "Fiber Optic Displacement Sensor With Subangstrom Resolution", Applied Optics, Vol. 29, No. 1, PP. 16-18, 1990.

5. D. Rugar, H. J. Mamin and P. Guethner, "Improved Fiver-Optic Interferometer For Atomic Force Microscopy", Applied Phys. Lett., Vol. 55, No. 25, PP. 2588-2590, 1989.

6. H. Stark, F. B. Tuteur and J. B. Anderson, Modern Electrical Communications: Analog, Digital and Optical, 2nd Ed., Prentice Hall, Englewood Cliffs, NJ, 1988, PP 313-314.

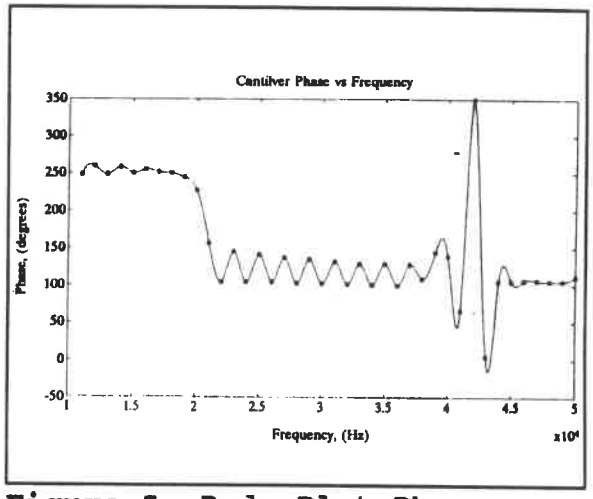


Figure 5. Bode Plot-Phase

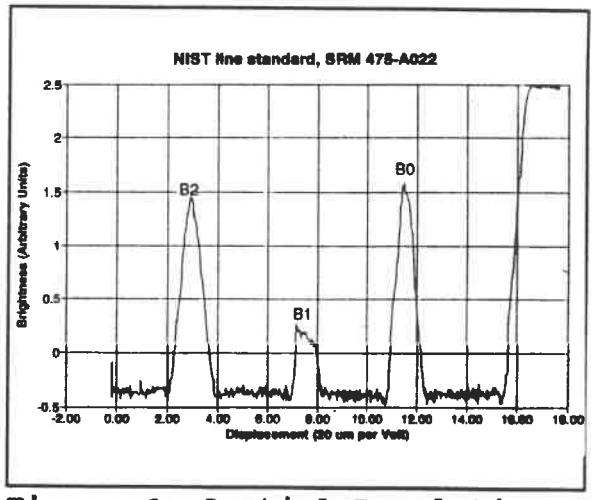


Figure 6. Spatial Resolution

Some observations on the performance of commercial phase measuring (Fizeau) interferometers used in surface figure metrology

Chris J. Evans and W. Tyler Estler

National Institute of Standards and Technology
Precision Engineering Division
Gaithersburg, MD 20899

There are a number of technological demands for aspheric optics with figure accuracies that challenge the state of the art in optical metrology. At the National Institute of Standards and Technology (NIST), we are in the initial phases of a five-year program aimed at improving the accuracy of optical figure measurements as a part of an Institute Competence Program in support of X-ray projection lithography. A number stepper designs for lithography at 13 nm have been proposed; many call for off-axis aspheric optics with figure accuracies of $\lambda/200$ or better (where $\lambda = 633$ nm). This far exceeds NIST's calibration capabilities for flats, let alone aspheres.

Production of optics to the tolerances indicated above will require close coupling of fabrication and metrology. Commercial phase measuring interferometers (PMI) are commonly used; thus the initial focus of our program is to assess the ultimate accuracy achievable with such instruments. Our long range goal is to provide high-accuracy calibration services for PMI reference optics to be used by optical manufacturers for in-process quality control.

Our first technical goal is to demonstrate the calibration of an optical flat (150 mm diam.) with an absolute accuracy of $\lambda/200$ or better. We have designed, and are in the final phases of construction, of a system (Figure 1) which will allow intercomparison of different approaches to absolute testing of flats. PMI measurements will be referenced to a fluid surface with diameter large enough to avoid meniscus errors. The same flat will be measured using multi-setup absolute calibration procedures such as the optical analog of the classic Whitworth three-flat test and the methods due to Parks and Fritz. The three flat test gives data only along a single diameter; calibration of the full aperture then requires a number of further, rotated set-ups; clearly, the spatial wavelengths of measured errors are constrained by the number of rotations performed.

Parks' method is based on the idea that the Zernike polynomials fitted to the measured optical path difference may be considered as the sum of Zernikes associated with the interferometer and those associated with the part. Thus if data is taken with the part rotated to a number of positions

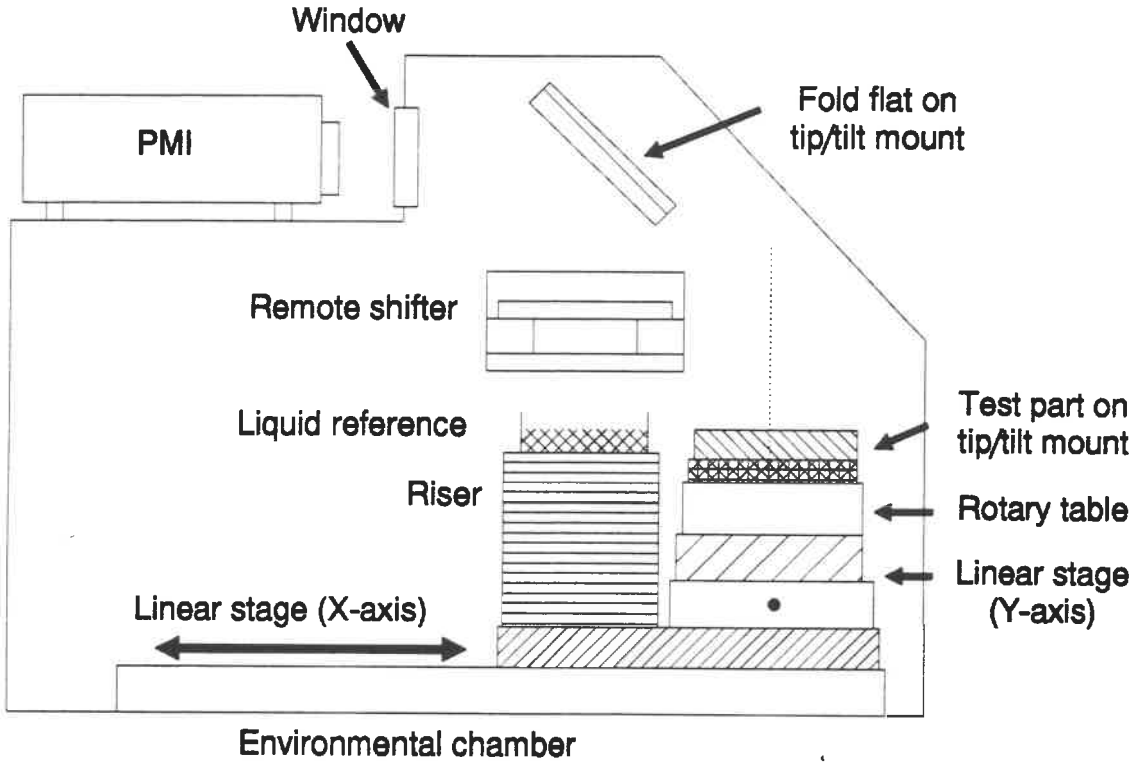


Figure 1 Schematic of flat testing system

about the optical axis, the terms associated with the part will rotate, and thus non-rotationally symmetric errors can be separated. Fritz combined these approaches, using three flats plus one rotation. Here, as with Parks' method, the spatial wavelengths of errors measured are limited by the number of Zernike terms.

The final approach to be evaluated is based on a new algorithm, developed independently at NIST and by Ai et al, using mechanical shearing of the cavity by an integral number of pixels in two orthogonal directions. It can be shown, mathematically, that this provides enough information to extract part and interferometer errors exactly over the common area to all three interferograms. In practice, however, the method will be limited to higher order aberrations since, for good optics, the slope is such that the difference between adjacent pixels is less than the resolution of current commercial PMIs. In addition, Ai has shown that removal tilts introduced in mechanical shearing leads to loss of curvature information.

In a later evolution of our program, spheres and mild aspheres will be measured with reference to ideal parabolae generated by ultrastable rotation of the fluid mirror.

Modern commercial PMIs typically have resolution of λ/1000 and, with reasonable care, can give excellent repeatability. To achieve absolute accuracy commensurate with this resolution and repeatability, we must understand the systematic error sources and the convolution of the part with the instrument. For example, we have observed a significant system sensitivity to fringe density. Figures 2 & 3 show the results of experiments in which the same flat was tested with varying fringe densities. Peak-to-valley (P-V) and RMS surface errors, with tilt removed the analysis software, is both observed to increase significantly with fringe density. The errors are not symmetric with respect to rotation about the cavity optical axis. Similar results are obtained when a sphere is tested with varying power.

The measured (P-V) figure error is also sensitive to the averaging procedure used when attempting to minimize the effects of air turbulence. For example, if one measures the sample P-V error four times and computes an average, the result is in general larger than when the PMI system is asked to take a series of four measurements and averages the data in the analysis software. A simple model, assuming a Gaussian noise distribution for each detector pixel in the imaging CCD array, provides a plausible explanation of this behavior and serves as a guide for efficient PMI data acquisition.

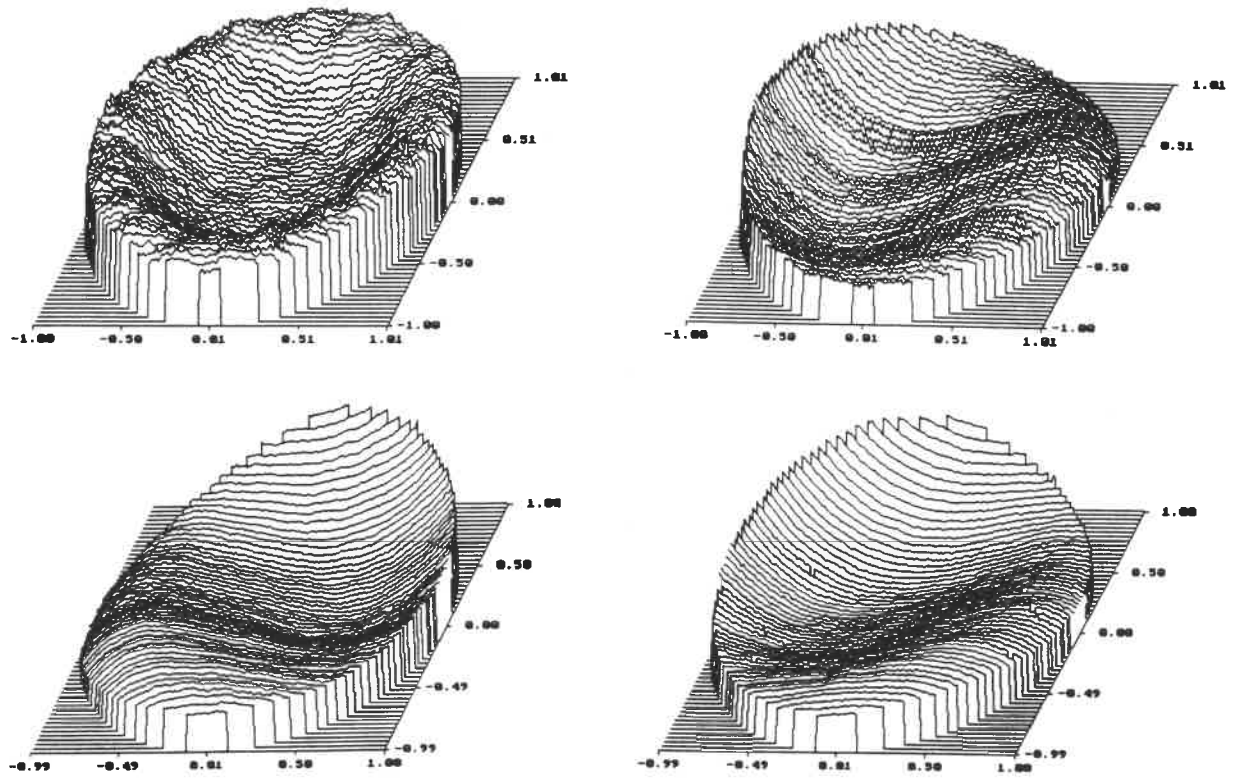


Figure 2 Three measurements of the same 150 mm dia flat: Top left - with a null fringe [P-V = 0.059 waves]; Remainder - 15 waves of tilt in different directions [P-V = 0.263, 0.291 & 0.34 waves]

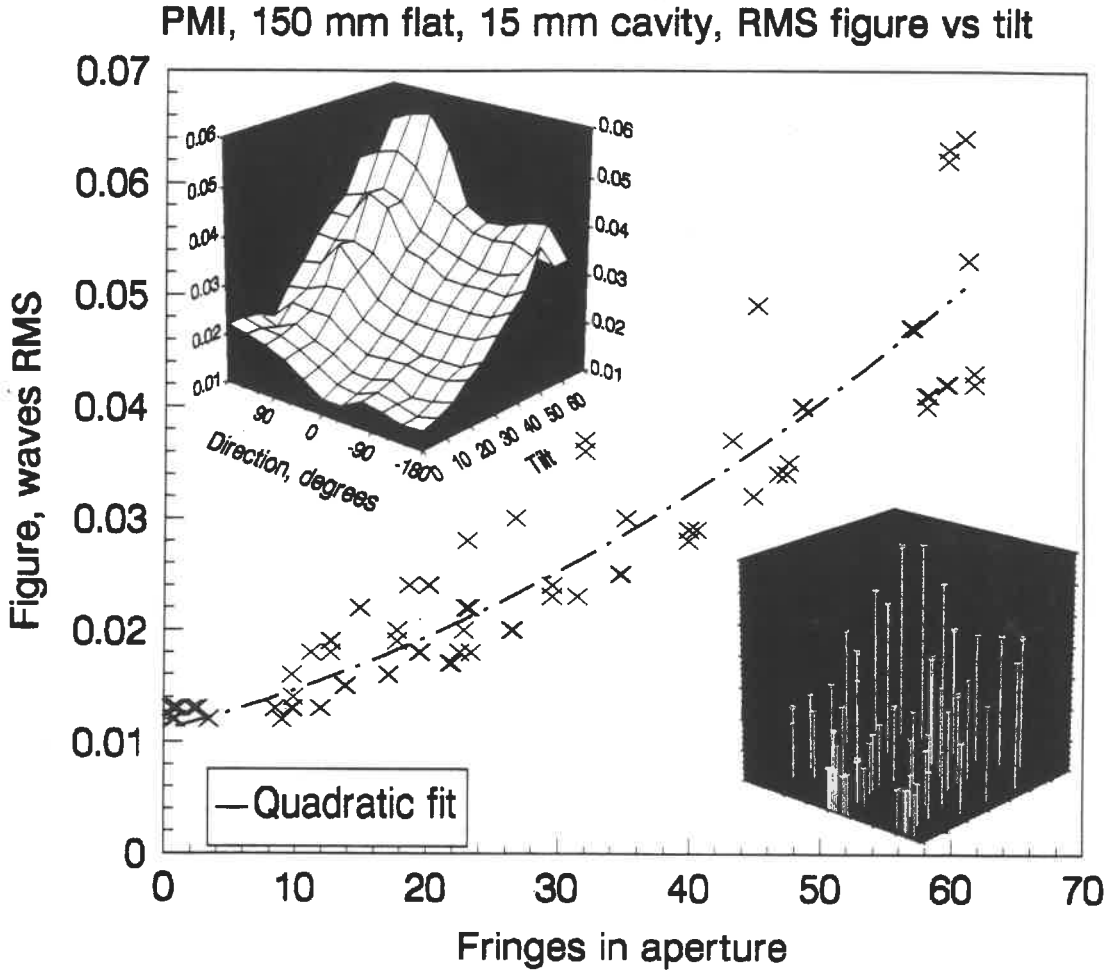

PMI, 150 mm flat, 15 mm cavity, RMS figure vs tilt
Figure, waves RMS
Fringes in aperture
Direction, degrees
Tilt
—Quadratic fit
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50
60
70
90
0
-90
-180
10 20 30 40 50 60

Figure 3

A FOCUSED VIEW ON ASPHERICS

By David M.G. Stevens

RANK TAYLOR HOBSON LTD.
P.O. Box 36, 2 New Star Road
Leicester LE4 7JQ England

INTRODUCTION

Modern aspheric forms are used to correct spherical aberration, in optical systems, which are used in every day items such as ophthalmic contact lenses and compact disc laser focusing systems.

Today the ratio of spherical departure and spherical radius is getting smaller and smaller in order to reduce the optical system dimensions, weight and increase the angle of field.

This no longer can be reliably produced by a polishing technique. An inexpensive method of manufacture had to be found, utilising less labour intensive techniques. Initially a grinding wheel was dressed with a diamond tool, this in turn was used to produce the desired aspheric form onto the workpiece. The early shapes used were of conic section profile (Ellipse, Parabola or Hyperbola).

It can be seen from Figure 1 that as the tool moves laterally the pivoted fulcrum arrangement produces a tip locus that scribes the desired shape.

By guiding the path of the machine tool with a steel band, which is formed around an involute shape, the desired surface form is produced. (Figure 2)

This method has an inherent characteristic that mechanically suppresses any small form errors from being transferred onto the component from the involute shaped centre.
The modern method of Aspheric surface generation is by precision single point diamond turning method or ultra precision grinding. Coupled with precision manufacturing techniques is the ability to measure the resultant surfaces to a higher accuracy.

METHOD OF ANALYSIS

The advent of the computer has enabled many analytical methods to be applied to a multitude of processes. Early metrology products had mechanical set-up facilities such as tilt adjustment, centring and levelling with associated analogue signal processing (high and low pass filtering).

Computerisation facilitates mathematical data manipulation previously not possible. The basic least squares straight line, for example, can be applied to a linear 2D signal to remove any surface slope caused by setup error.

A least squares circle or partial arc can be determined by mathematically varying the radius value and the X/Y coordinates of the radius centre position to derive the optimum circle fit.

With each of the straight and curved line methods, the geometric elements (slope, offset, radius and centre) are individually varied to optimise a fit. Alternatively, absolute data can be used in a least squares algorithm to determine the deviation from the ideal form.

This is ideally suited to the measurement of aspheric forms where the absolute shape is described by a polynomial expansion formula.

METHOD OF MEASUREMENT

The aspheric form is measured by a contacting stylus instrument, relative to a straight line reference datum bar, to a resolution of 10nm in the 'Z' axis with a maximum range of 6.0mm and 0.25µm in the 'X' axis with a maximum range of 120mm.

The dynamic range must be continuous (i.e. without magnification change) which demands a specific type of gauging system. The conventional inductive gauging system works on the principle of a transformer, with a variable amount of primary voltage being induced into a secondary coil in a pivoted arm arrangement (fig. 3). This method of gauging requires modulation, demodulation, signal amplification, digitisation and analysis. Typically the digitisation is performed with a 15 bit analogue to digital convertor, which produces a 32768:1 range to resolution ratio. This is insufficient for practical, simultaneous, form and surface texture analysis to the required accuracy demands.

For this reason a Michelson Interferometer is employed as part of a contacting gauging system, which has a continuous range to resolution ratio of 600,000:1. This enables simultaneous form and surface texture measurement over a dynamic range of 6.0mm to a resolution of 10nm.

The system is capable of detecting a stylus movement of $\pm$ 3mm in the 'Z' axis, derived from a change in laser light path length having passed through a cube corner prism attached to the rear of a pivoted stylus arm (fig.4). The cubed corner giving a constant internal optical path length which is not influenced by lateral position and/or orientation.

This changing path length when compared to the reference source, changes phase producing an interference fringe pattern which is proportional to the path length change. These fringes occur every half wavelength, using a Helium/Neon laser but due to the double folded path this equates to 158.2 nm.. These fringes are detected by suitable photoelectric detectors and then interpolated by a factor of 16 to obtain the required resolution of 10nm, the interpolation is derived from the amplitude relationship of these quadrature signals.

The mechanical arrangement of the pivoted stylus is an arcuate movement, which introduces positional error in both 'Z' and 'X' directions. Compensation is made by the microprocessor so as to follow the true locus of the stylus tip.

Using a precision ball, of known radius and of suitable surface texture, the only requirements are to crown the ball in both 'X' and 'Y' axes and enter the absolute ball and stylus tip radii. The system then enters an automatic sequence and derives the four factors from the measured data.

These constants give first and second order compensation in both 'Z' and 'X' axes which are proportional to the perceived 'Z' axis position:

$Z' = Z*A + Z^2*B$ & $\Delta X = Z*C + Z^2*D$

$X' = X - \Delta X$ where Z is the perceived Z' axis position.

The mechanical set-up of the aspheric component to be measured is essential. Levelling in the 'X' axis is not critical as the least squares algorithm will adjust accordingly, however, the 'Y' levelling is critical as no compensation can be mathematically applied.

The stored data is then subjected to a least squares fit using the polynomial expansion discussed earlier. The residual data is then shown as a difference graph, relative to a straight line representing the ideal aspheric form.

The manufacturing process has two main areas of error:-

1. Absolute radius error, usually due to tool tip dimension.
2. Shrinkage, in the case of optical component manufactured from an aspheric master mould (i.e. soft contact lenses).

This residual error can result with a concave or convex residual error being exhibited. (See fig.5) If over compensation is made to the radius term the residual error can invert (see fig.6). Until an optimum value is found where the residual error becomes a minimum. (See fig. 7) By observing the radial error (sag), compensation can be made in the manufacturing cell to correct the source of error (tool tip radius or setting).

REMARKS

When a specimen is analysed with an interferogram technique, the component appears to be perfect. This is because the polynomial departures from the base shape is retained as this differential information is correct. However, no absolute information can be obtained from the differential comparator, therefore large radial error can exist and remain undetected until later in the assembly process.

The method of measurement described in this paper is an absolute measurement technique which does not suffer from deficiency that the Interferogram exhibits.

Ultimately, a method of testing whilst in-process is desirable. This method would probably have to be non-contact, due to mechanical frequency response of the stylus assembly and the surface measurement speed. However, any machine motion error would also be compoundly measured, which would appear different on the static component.

ACKNOWLEDGEMENTS

The Author wishes to thank the Directors of Rank Taylor Hobson Ltd. for their permission to publish.

REFERENCE

1. Mr. J.D. Garratt, Application for a wide Range Stylus Instrument in Surface Metrology - Second International Conference on Metrology and Properties of Engineering Surfaces.

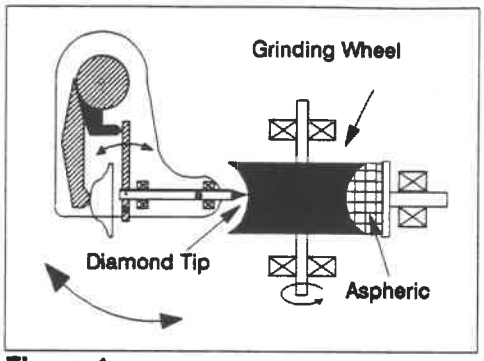


Figure 1

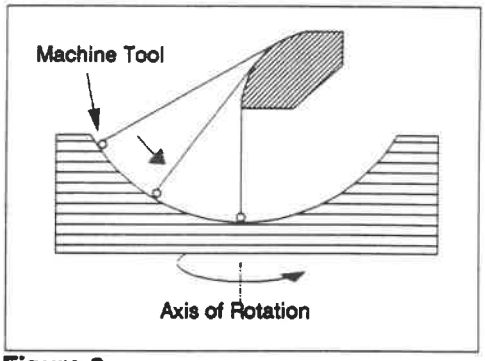


Figure 2

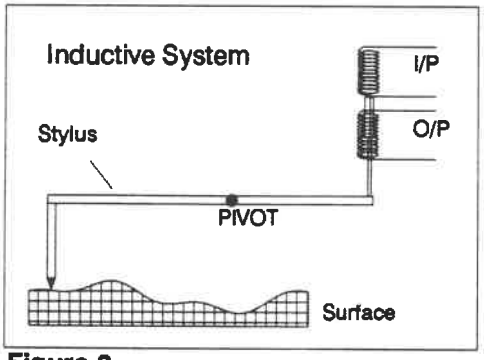


Figure 3

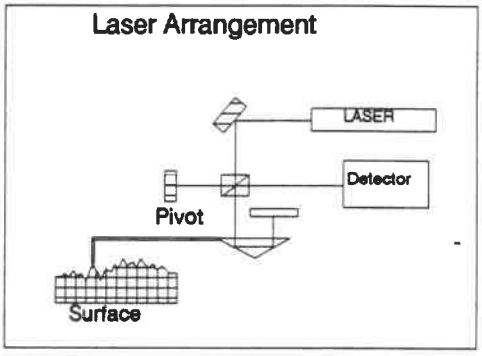


Figure 4

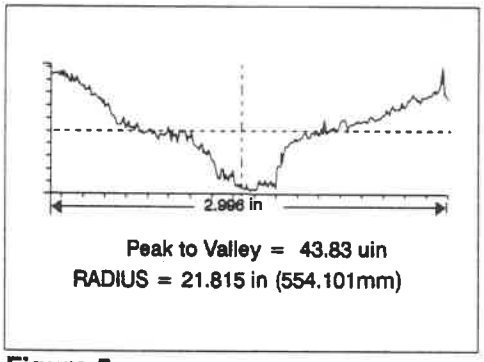


Figure 5

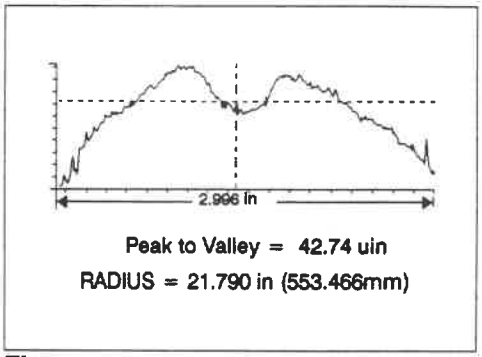


Figure 6

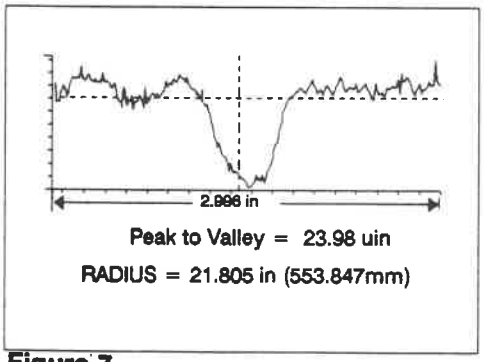


Figure 7

CONTROLLER DESIGN FOR A MODERN DIAMOND TURNING MACHINE

Michele H. Miller
Kenneth P. Garrard
Thomas A. Dow
Lauren W. Taylor
Precision Engineering Center
North Carolina State University
Raleigh, NC 27695-7918

INTRODUCTION

The technique of machining optical quality finish on a diamond turning machine (DTM) has revolutionized the fabrication of reflective and transmissive optical components. This technique frees the optical designer from the constraints of spherical shapes and also provides a method of integrating the optical surface into the mechanical structure. By adding the possibility to fabricate non-rotationally symmetric surfaces with a fast tool servo (FTS), an even more dramatic expansion in the capability of optical components and reduction in their cost can be foreseen.

The limitations of the currently available commercial controllers revolve around the processors and the openness of the architecture. High-speed I/O is required for the large volume of data and the fast data rate necessary for a non-rotationally symmetric part. Current controllers are not fast enough to generate the reference positions, to accomplish the control algorithms necessary for the machine axes, or to provide correction of the geometric errors. In addition, the closed architecture does not allow additions of new inputs or outputs or major changes in the control algorithms.

A new controller design is proposed with the capability to extend the potential of the DTM. The architecture of this design is presented as well as a demonstration of the potential for shape generation and error correction.

CONTROLLER ARCHITECTURE

A series of controller architectures has been developed to improve the capability of DTMs. The first used a PC-AT/68020 computer system [1]. Important features were a 1 msec update time, real-time reference point generation capabilities for conic surfaces, and flexible control algorithm implementation. This controller, however, is limited in data transfer capability by the PC bus, and in computational power by the 68020 processor. While this design significantly improved the DTM, it could not be extended to more computationally intensive applications.

The solution to these problems led to the development of the H^2ART architecture. This heterogeneous, hierarchical architecture [1] utilizes a pair of DSPs (TI TMS320-C30) capable of rapid data passing and high-speed floating point calculations. DSP based controllers have been developed for fast tool servo applications which operate independently of the axes controller [2,3,4]. Independent FTS control has accuracy limitations, however, as the FTS motion cannot be perfectly synchronized with the axes motions. The H^2ART architecture overcame these difficulties but has required more hardware and software effort than originally expected.

As an intermediate step in the development process, an integrated (but add on) controller has been built for application to current DTMs. This design uses a single H^2ART node to control the high-speed fast tool servo axis with an interface to monitor the position of the DTM axes. This design will extend the capability of current DTMs to fabricate non-rotationally symmetric surfaces without requiring a complete control system change.

A new DTM controller incorporating the 68020 based axes controller and a H^2ART based FTS controller is depicted in Figure 1. The 68020 controller has the processing speed to

implement modern control algorithms for the axes motion as well as slideway error corrections. The H²ART based controller with I/O interfaces to the DTM allows integration of the FTS control with the axes control.

Interface Boards

Three logical interfaces are needed to build an integrated controller that combines the H²ART architecture, FTS actuator, Pneumo ASG 2500 DTM, Zygo Laser Interferometer, and 68020 controller components. These three interfaces are described below.

Angular Feedback Interface This interface provides the FTS controller with the current value of θ expressed as the number of encoder pulses since the last once-per-revolution index pulse. The maximum resolution of the encoder is 0.0015 radians per count. The direction of rotation can also be determined. The once-per-revolution index pulse is available externally to facilitate angular alignment of the workpiece fixture.

Laser Input Interface This interface maps the bits of the 50-pin interferometer interface connectors (one per axis) into one 8-bit control input port (both axes), one 8-bit control output port (both axes) and two 16-bit data input ports (one per axis). This gives the FTS controller timely access to the current X and Z axes positions.

Position Feedback Interface The 68020 controller also needs access to X and Z. However, it is not feasible to connect two controllers directly to the interferometer interface output ports, since they would both attempt to drive the measurement board sampling signals. The Position Feedback Interface replicates the signals that the 68020 controller expects during communication cycles with the interferometer measurement boards. It also provides one level of data buffering. This permits the FTS controller to communicate current axis position information to the 68020 controller without synchronization delays. The only requirement is that the FTS controller supply current X and Z values to the interface board at a higher frequency than the 68020 controller update cycle (i.e., greater than 1 kHz). This is depicted in Figure 1 as switch position *a*. The Zygo Interferometer can be manually reconnected to the 68020 controller (switch position *b*) when the FTS is not in use.

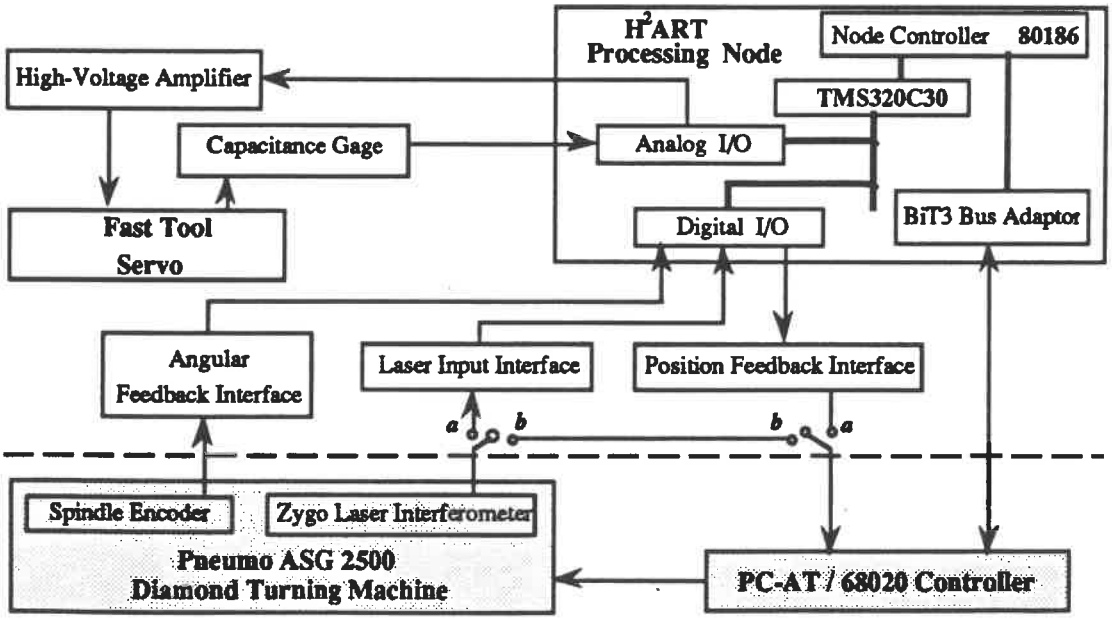


Figure 1: Integrated FTS controller configuration

CONTROLLER APPLICATIONS

The controller design has been used for two applications requiring high-speed communication and integration of the axes motion. The first is to fabricate non-rotationally symmetric surfaces and the second is to correct for the geometric errors in the motion of the machine slideways.

FTS Integration

Non-rotationally symmetric surfaces can be machined using a high-bandwidth servo actuator to generate small perturbation Z-axis motions synchronized to the rotation of the workpiece spindle (θ) and the radius (X). The remainder (or base) surface is described by a best-fit aspheric surface of revolution. Being independent of θ, this base surface can be machined using the standard control system on the DTM. The motion of the tool is corrected for the non-rotationally symmetric part of the shape by the FTS. The FTS is driven by a high-voltage amplifier directly from the H^2ART analog I/O board. Servo position feedback via a capacitance gage is also interfaced through this analog board. The FTS feedback control loop is independent of the DTM slide control. However, generation of the non-symmetric reference signal requires access to the spindle encoder and the Zygo interferometer.

Reference Surface Generation The decomposition of the general aspheric optics equation into (X)-dependent and (X,θ)-dependent components suitable for on-axis FTS machining of off-axis conic segments is described in [3,5]. This solution minimizes the range of motion required of the FTS by tilting the workpiece. Parametric forms of the conic surfaces [1] are being investigated as a means of improving the accuracy of the best-fit asphere by producing equal arc length increments on the surface. The best fit asphere is generated off-line and loaded into the 68020 memory, while the non-rotationally symmetric correction is made by the H^2ART node with respect to the actual position of the slides. (A fast processor, i.e., another H^2ART node, could also be used to generate the asphere *on-the-fly*.)

The integrated FTS controller calculates the Z-axis reference position (as a function of X and θ) *on-the-fly* and subtracts the current Z axis position to obtain the small perturbation component. An example of the non-rotationally symmetric component of an off-axis parabola is shown in Figure 2. This surface is a 5 in diameter section of a 42.5 in focal length parabola 11.5 in from the center. The range of the servo is only 600 µin.

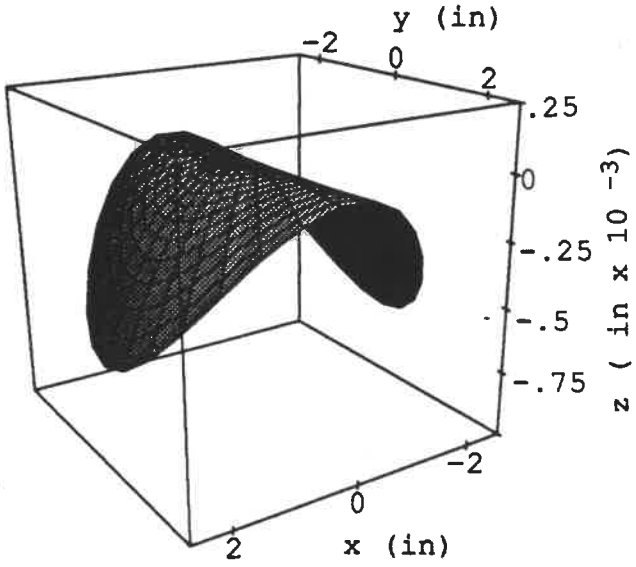


Figure 2: Non-rotationally symmetric component of an off-axis parabolic mirror

By using the actual axes positions in the reference calculation, the FTS can also compensate for any high frequency Z axis vibrations as well as the low frequency positioning error of the 68020 axes controller.

Slideway Error Compensation

Non-repeatable axes errors have been measured and compensated in real-time [6]. If measurements of axes errors show that they are repeatable, a permanent error map can be generated and used to modify X and Z reference points off-line on a part tape or in real-time during reference point calculation.

To implement a real-time straightness error compensation algorithm, an error table was created for each axis based on error measurements. Each table contains the axis position and the straightness error (or the

amount of correction to be applied by the other axis). A fixed error interval of 2.5 nm (1 laser count) was chosen to separate table entries. Despite the small error resolution, the error tables for both slides are short--156 entries for the X slide and 79 entries for the Z slide. Each row of an error table contains a cumulative axis error value and the position of the other axis at which the error is applicable. Linear searches through the error tables are avoided by noting that on each controller cycle the array index of the correct cumulative error is at most one table entry away from its position on the previous cycle. Thus, the geometric error compensation procedure is independent of the size of the error tables and can be calculated in constant time. As shown in Figure 3, it is executed after the reference coordinates for each slide are calculated but outside the feedback loop.

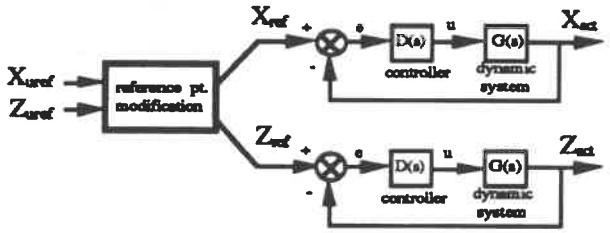


Figure 3: Block diagram which includes slideway error correction

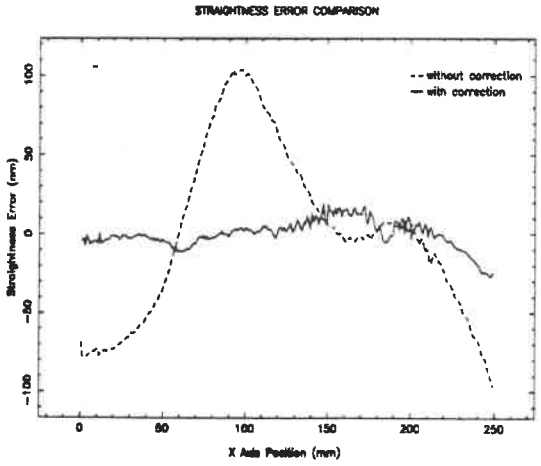


Figure 4: X straightness measurement with and without controller compensation

To test the algorithm, axis straightness was measured with and without compensation. Figure 4 shows the results for the X slide. Correction reduces the error from 200 nm PV to 40 nm PV. The correction technique was also applied to the Z slide but because the error was only 60 nm, the results were less dramatic. The error was reduced to 40 nm PV.which is the repeatability limit of the straightness measurements.

CONCLUSIONS

Ultimately a modern DTM controller will use fast DSP's for all of its control processes. A fast processor for X and Z control, for instance, could execute real-time tool radius corrections and higher order control filters. A modern controller should also be capable of transferring previously calculated FTS reference points from storage to the FTS control processor. With the current configuration they must be calculated *on-the-fly*. A H^2ART architecture which incorporates two DSP's (thus eliminating the 68020 processor) can achieve these requirements and is the next step. However, the intermediate controller configuration described here achieves FTS integration and has the advantage of easy portability to commercial axes controllers.

REFERENCES

[1] Fornaro, R. J.; Garrard, K. P., Taylor, L. W., "Architecture and Algorithms for Computer Control of High Precision Machine Tools," *Proceedings of ASPE Annual Conference*, 1990.

[2] Dow, T. A., Miller, M. H., Falter, P. J., "Application of a Fast Tool Servo for Diamond Turning of Non-Rotationally Symmetric Surfaces," to be published in *Precision Engineering*.

[3] Luttrell, D. E., "Machining Non-Axisymmetric Optics," *Proceedings of the ASPE Annual Meeting*, 1990.

[4] Weck, M., Pyra, M., "Diamond Turning of Non-Symmetric Laser Optics," *Proceedings of the ASPE Annual Meeting*, 1990.

[5] Gerchman, M. C., "A Description of Off-Axis Conic Surfaces for Non-Axisymmetric Generation," *SPIE Proceedings*, Vol 1266, 1990.

[6] Donmez, M. A., et. al., "A general methodology for machine tool accuracy enhancement by error compensation," *Precision Engineering*, Vol. 8, No. 4, 1986, pp. 187-196.

Nanometer Motion Control via Magnetic Suspension

David L. Trumper
Dept. of Electrical Engineering
Univ. of N. Carolina at Charlotte
Charlotte, NC 28223

Alexander H. Slocum
Mechanical Engineering Dept.
Massachusetts Institute of Technology
Cambridge, MA 02139

Magnetic bearings represent a promising approach for achieving positioning with nanometer resolution. The authors have constructed a high-precision linear bearing which demonstrates 5 nm short-term position stability. This system represents the experimental implementation of an idea presented in [1] and has been investigated as part of the first author's Ph.D. thesis [2].

This study was initiated with support from the Molecular Measuring Machine (M^3) project at the National Institute of Standards and Technology [3]. The form-factor of the bearing was thus chosen such that it was compatible with the M^3 crossed-slide design. The bearing technology is also highly suitable for other precision applications, such as in stages for photolithography and diamond turning machines; these application areas provide the focus for our current research.

As shown in Figs. 1–4, a 10.7 kg platen measuring 125 mm by 125 mm by 350 mm is suspended and controlled in five degrees of freedom by seven electromagnets, labelled as $B1$ through $B7$. Five capacitive probes located in the bearing centers measure position with nanometer resolution. The seven points at which the bearing forces act on the platen are shown as arrows and the five points at which the position measurements are taken are indicated as dots in Fig. 1. The scale of Figs. 2, 3, and 4 is indicated by the 100 mm reference shown in Fig. 2.

The suspension acts as a linear bearing, allowing linear travel of 50 mm in the long axis of the platen. In the laboratory this bearing system has demonstrated position stability of 5 nm peak-to-peak. This is believed to be the highest position stability yet demonstrated in a magnetic suspension system. Performance at this level confirms that magnetic suspensions can address motion control requirements at the nanometer level.

In previous work [2,4,5], the long-travel degree of freedom is not actively controlled. Efforts are currently directed at achieving control of this degree of freedom with 50 mm travel through a novel magnetic suspension linear motor. This motor consists of an array of permanent magnets attached to the moving platen and a stator array of coils attached to the fixed machine frame as shown in Figs. 2–4. Prototypes of this linear motor are currently under construction. Position in this degree of freedom is transduced by a laser interferometer. Optics for this interferometer are not shown in the figures.

The key feature of the linear motor is that it affords control of forces in two degrees of freedom, one parallel to the air gap and lying on the axis of the magnet array and one perpendicular to the air gap. Thus it is possible to simultaneously control translation and suspension with a single motor structure.

To calculate the electromagnetic forces, the motor is idealized as a two dimensional structure as shown in Figure 5. Here, the motor is assumed to have a depth of w into the paper and to extend indefinitely in the $\pm z$-direction. The stator is fixed in the laboratory frame x, y, z. The primed coordinate frame x', y', z' is fixed in the layer of magnetization, and is displaced from the unprimed frame by $(x_0 + \Delta)\vec{i_x} + z_0\vec{i_z}$. The stator is modeled by a sinusoidally distributed, y-directed current density $J_y = \mathrm{Re}\tilde{J}e^{-jkz}$, in a layer from $x = 0$ to $x = \Delta$. The complex amplitude $\tilde{J}$ is a function of the winding density η_0, and the two phase currents, I_1 and I_2. The magnet array is modeled by a half-infinite region of sinusoidally distributed, x-directed magnetization $M_x = \mathrm{Re}M_0e^{-jkz'}$. The magnetization layer is referred back to the unprimed frame by substituting $z' = z - z_0$ into M_x which gives $M_x = \mathrm{Re}\tilde{M}e^{-jkz}$, with complex amplitude

$\tilde{M} = M_0 e^{jkz_0}$. Note that the complex amplitudes represent the temporal variations in both amplitude and spatial phase.

For this structure, the forces acting on an integer spatial period length l of the magnet array are given by

$$\begin{bmatrix} F_x \\ F_z \end{bmatrix} = \eta_0 M_0 G e^{-kx_0} \begin{bmatrix} -sin(kz_0) & cos(kz_0) \\ cos(kz_0) & sin(kz_0) \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (1)$$

where $G = \frac{\mu_0 wl}{4k}(e^{-k\Delta} - 1)$.

In operation, the controller will invert the force equations given above in order to supply commutation to the motor. That is, given desired forces in the x and z directions, F_{xd} and F_{zd}, respectively, the currents necessary to achieve these forces are given by

$$\begin{bmatrix} I_{1s} \\ I_{2s} \end{bmatrix} = \frac{e^{kx_0}}{\eta_0 M_0 G} \begin{bmatrix} -sin(kz_0) & cos(kz_0) \\ cos(kz_0) & sin(kz_0) \end{bmatrix} \begin{bmatrix} F_{xd} \\ F_{zd} \end{bmatrix}. \quad (2)$$

where I_{1s} and I_{2s} are setpoints to the two-phase current drive which sets I_1 and I_2.

Given such a linear motor, various stage geometries are possible. One such geometry is shown in Figure 6. Here, four linear suspension motors are combined to act on the top surface of a platen to implement an X–Y stage suitable for photolithography. The stators are labelled $S1$ through $S4$ and drive against four magnet arrays which are shown as grey rectangles. Stators $S1$ and $S3$ drive the platen in the Y-direction and stators $S2$ and $S4$ drive the platen in the X-direction. The four stators are driven appropriately to control the remaining four platen degrees of freedom.

The Linear Modal Compensator block consists of separate lead/lag compensators for each of the six degrees of freedom. The Compensate Geometric Nonlinearities block translates forces and torques specified in the platen frame (where the system appears as linear) into forces specified in the stator frame (where they are actually applied to the platen). The Motor Commutation and Power Amp blocks each implement (2) and contain power amplifiers for each of the stator phases.

The linear bearing system of Figs. 2–4 provides a test-bed for a single linear suspension motor. The design will be refined here before advancing to the X-Y system shown in Fig. 6.

References

[1] Slocum, A.H., and Eisenhaure, D.B., "Design Considerations for Ultra-Precision Magnetic Bearing Supported Slides," NASA Magnetic Suspension Technology Conference, Hampton, Va. Feb. 2–4, 1988.

[2] Trumper, D.L., "Magnetic Suspension Techniques for Precision Motion Control," Ph.D. Thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., Camb., Mass., Sept., 1990.

[3] Teague, E.C., "The National Institute of Standards and Technology molecular measuring machine project: Metrology and precision engineering design," *Journal of Vacuum Science Technology*, B7 (6), Nov/Dec, 1989, pp. 1898-1902.

[4] Trumper, D.L., and Slocum, A. H., "Five-Degree-of-Freedom Control of an Ultra-Precision Magnetically-Suspended Linear Bearing," NASA Workshop on Aerospace Applications of Magnetic Suspension Technology, NASA Langley Research Center, Hampton, VA, Sept. 25–27, 1990.

[5] Trumper, D.L., "Precision Magnetic Suspension Linear Bearing," International Symposium on Magnetic Suspension Technology, NASA Langley Research Center, Hampton, VA, Aug. 19–23, 1991.

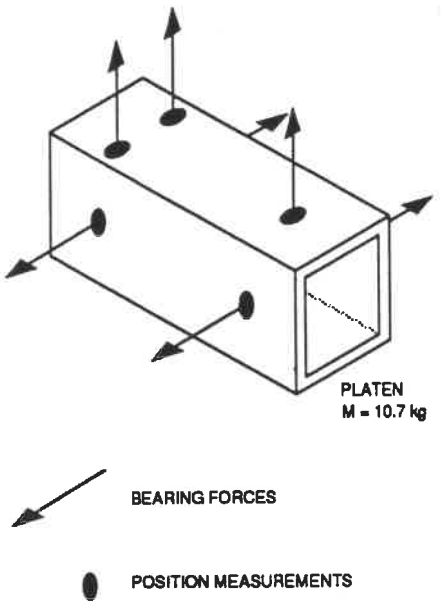


Figure 1: Bearing Forces

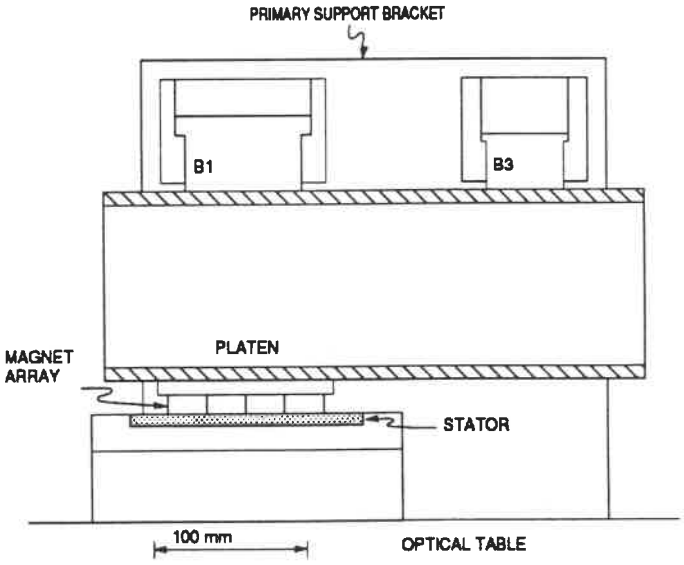


Figure 2: Bearing Assembly Side View

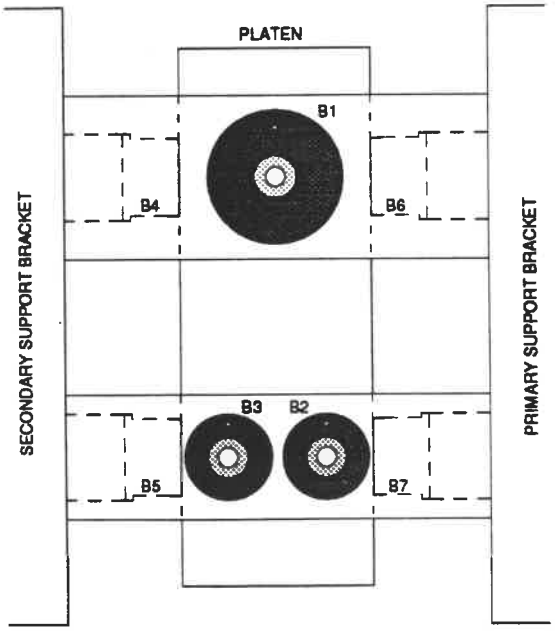


Figure 3: Bearing Assembly Top View

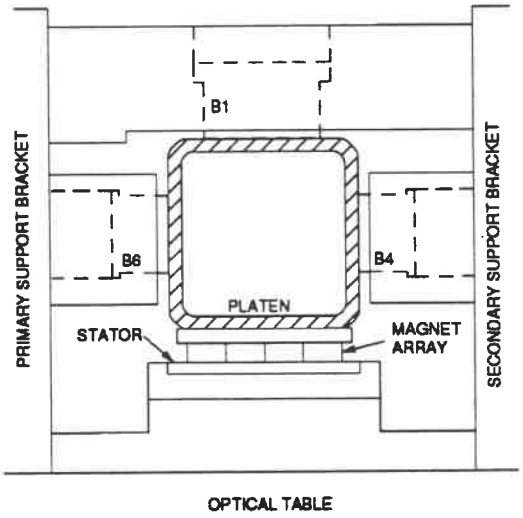


Figure 4: Bearing Assembly End View

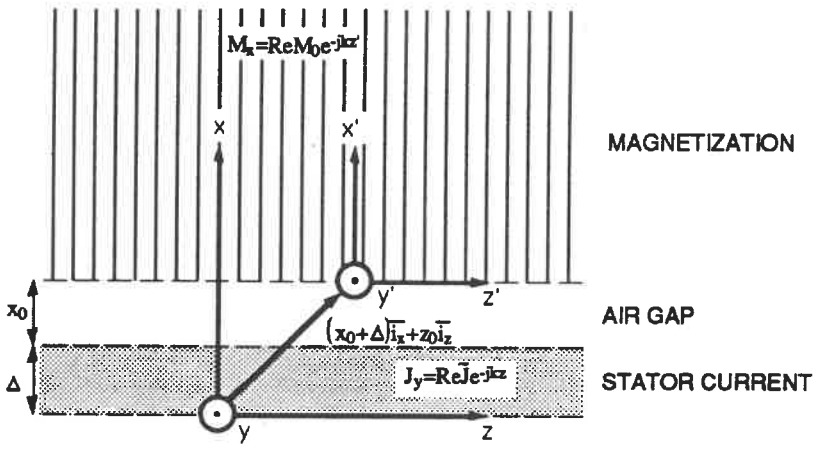


Figure 5: Motor Model

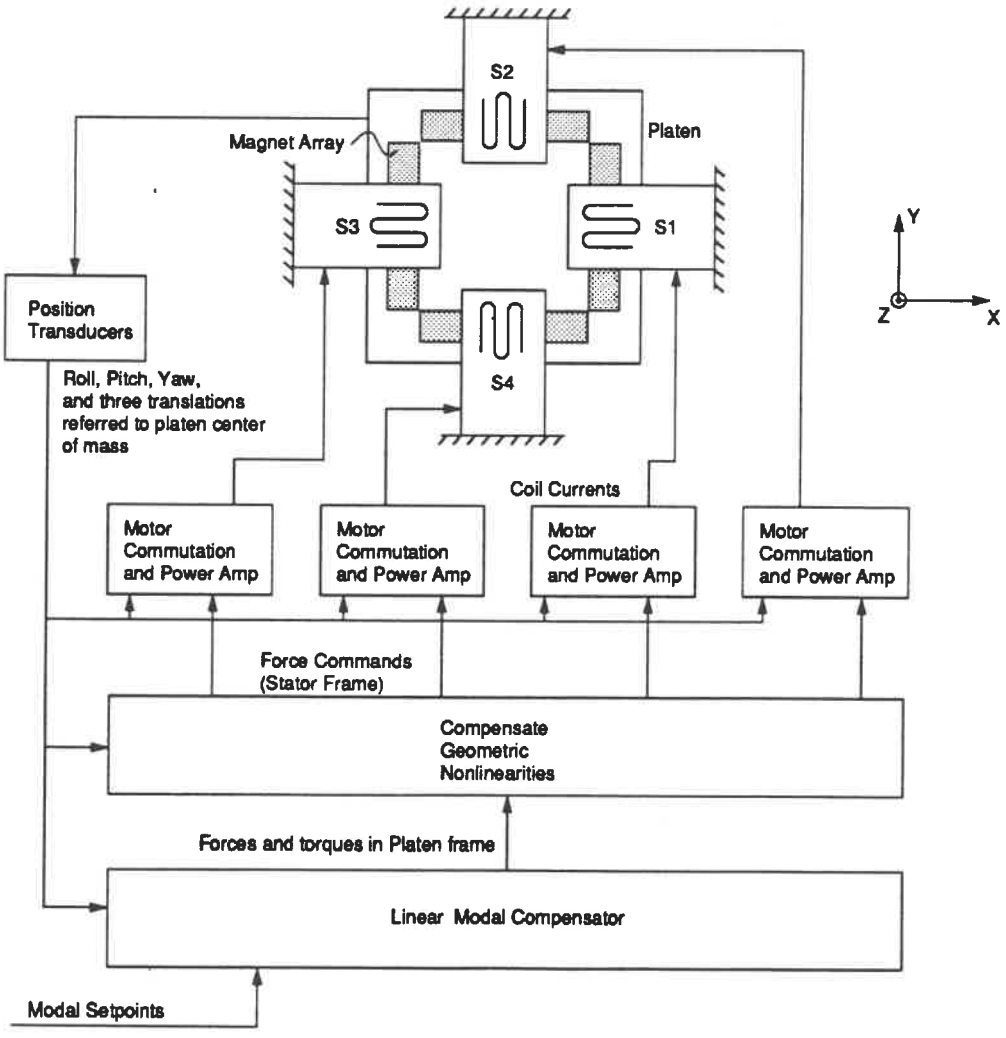


Figure 6: X-Y Stage Controller Block Diagram

"DIAMOND POINT MACHINED SNAP-TOGETHER OPTICS"

John P. Schaefer
Texas Instruments
Dallas, TX

ABSTRACT

A "snap-together" optical assembly consists of any two or more diamond point machined components which achieve optical alignment when mechanically assembled. This paper describes the design and manufacture of snap-together diamond point machined optics for infrared applications. Design issues such as stress isolation, alignment tolerances, and producibility issues that effect quality, cost, and performance are discussed. Assembly techniques are explained and several snap-together assemblies that demonstrate sound design practices are shown. A properly designed and manufactured snap-together system can be extremely producible, cost affective, and essential in pushing the state-of-the-art in IR optical system design and performance.

INTRODUCTION

The ability to cost effectively manufacture reflective mirrors out of lightweight non-ferrous materials has given optical designers a tremendous amount of design freedom. System designs which incorporate diamond machining technology are more reliable and have higher performance. Tolerance stack-ups are reduced when the mirror, mirror mount, locating surface, and housing are integrated into a single unit and machined in the same set-up. By producing the mating parts in the same manner, common alignment errors which affect optical performance are reduced. The benefits of a properly designed snap-together system include; reduced components, greater structural integrity, reduced assembly and testing time, and improved reliability and performance. All of these benefits lead to an overall reduced system cost.

DESIGN CONSIDERATIONS

An important consideration when designing a DPT element is proper mounting stress isolation. Mounting strains which propagate to the optical surface can create severe surface irregularities. A kinematic three-point mounting scheme is the most common and effective design for reducing mounting strains and "springing". A very good example is shown in the optical system pictured in Figure 1. It can be seen that the large webbed off-axis parabola is bolted to it's backing plate in three areas. The plate is, in turn, bolted to the housing and also uses a kinematic three-point mount. Two sets of locating pins on the backing plate control rotation and position of the parabola. Backing plates are inexpensive to produce and effectively isolate the mirror from the main structure.

The opto-mechanical designer needs to determine the degree of stiffness and compliance needed for the mirror to be securely mounted with minimal distortion. The most common method used to produce compliant mounting pads is to weaken the pads by using either circular or horizontal saw cuts.

These cuts are necessary for two reasons;

(1) so that mounting pads are compliant with the mating surface to minimize stress affects from flatness mismatches and,
(2) to prevent torsional mounting strains from propagating to the optical surface.

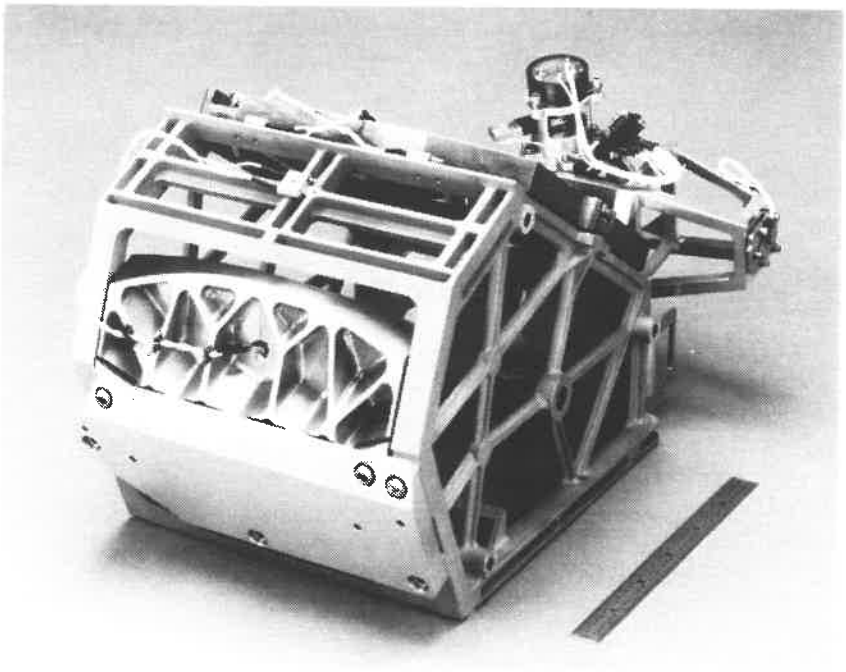


Figure 1. Snap-Together Missile Seeker Assembly

There is a fine line between stiffness and compliance of the mounting pads.

If the mounting isolation measures are too extreme, several problems can occur;

(1) the pads will not maintain their co-planarity during initial lathe turning operations and will require re-machining to achieve the desired pad flatness prior to final finishing,
(2) the lack of adequate stiffness will cause vibrations during the machining process, and
(3) centrifugal affects will be magnified.

All of these conditions can cause surface figure errors to be present even in the as-machined, constrained condition.

Whenever possible, 3D solid modeling should be used to determine the effectiveness of stress isolation. If solid modeling is not available, test prototypes with varying degrees of stiffness versus compliance can be machined prior to finalizing the design. One other option is for the prototype design to be overly stiff around the mounting pads. Deeper mill cuts can then be added during the prototype machining until the proper amount of stiffness verses compliance is achieved. Fortunately, most mounting problems associated with faulty designs can be identified during the DPT operation and the problems resolved prior to assembly and system testing.

The source of surface figure errors can usually be pinpointed by testing the optic in the constrained, and unconstrained condition. If the surface is irregular in the constrained condition, lack of adequate pad support and centrifugal effects are usually the cause and an over-compliant situation has occurred. If there is no surface irregularity present in the constrained condition, but present un-constrained, the pads are too stiff and causing the classic "springing" situation.

One of the main advantages of snap-together assemblies is reduced assembly time because of the elimination of tedious and time-consuming alignment procedures. In order for the optical assembly to perform at peak efficiency, the components must accurately locate in relation to one another in the assembly. There are many ways to achieve accurate snap-together optical alignment but it is essential for the opto-mechanical designer to work closely with the DPT engineer so that tooling concepts allow components to be machined with minimal set-ups.

Axial symmetric cassegrain type assemblies should be designed so that axis centering requirements are achieved by machining the outside locating diameter of the concave primary mirror in the same operation as its mirror surface is machined. The mating inside diameter of the convex secondary mirror is also machined in the same operation as its mirror surface. (See Figure 3) The optical axes of these two mirrors are accurately aligned when assembled. To control tilt, exterior mounting surfaces are machined parallel to each mirror surface during the same operation that the bores are machined. Spacing is controlled by the machine and the CNC program. Using this approach, the expected tilt and spacing tolerance between the two components is tied to the DPT lathe's positional accuracy, axis straightness and perpendicularity. Typical machine tool specifications are less than 20 micro-inches (.5 microns) axis straightness, less than 1 arc-sec perpendicularity of the axis over full travel, and positional accuracy of less than 1 microinch (.025 micron). Actual tolerances over smaller travel distances are even less.

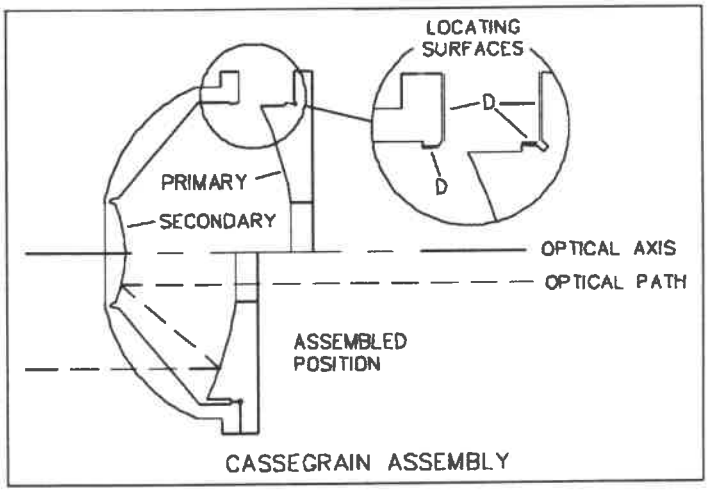


Figure 3

The degree of optical alignment achievable on off-axis assemblies is less than that of on-axis assemblies because of the unavoidable stack-up of tolerance between mating components. Off-axis beam expander type assemblies usually consist of at least five different components and require the use of locating pins to achieve optical alignments. These pins are used for rotational control and off-axis positioning. Focus, or spacing, is controlled by the flycut dimension on the housing, the center thickness of the parabolas and radius of the parabolas. Tilt is controlled by the parallelism of the flycut housing. Although off-axis beam expander assemblies contain several components and require the use of locating pins, snap-together optically aligned assemblies are producible in production quantities.

ASSEMBLY REQUIREMENTS

The designer must not only consider machining operations but must also be aware of assembly requirements. It is important that all mating surfaces have generous chamfers to reduce sharp edges and burring. Threaded holes should have generous countersinks to reduce thread pullout and aid the assembler in locating the tapped holes to eliminate surface scratching during assembly. Location pin holes should also be countersunk and should be below the diamond point machined mating surfaces.

Diamond pins should be used when tight rotational control is required. They should always be press fit into the mating surface and clocked. Diamond pins allow for radial distance tolerance stack-ups which can cause pin binding and a constrained condition. It is best to locate these pins in a counter bore so that material displaced during the pressing operation does not adversely affect the flatness of the locating surface.

A good rule of thumb to follow when determining the locating pin height above nesting surface is 1.5 times the diameter of the pin. This is all that is necessary to locate the components and reduces the chance for binding of the pin in the holes. This rule should also be followed when designing DPT lathe fixturing.

Proper torque procedures should be followed when assembling the components. The amount of torque the designer specifies should be adequate for the components to maintain alignment during and after system vibration and G-force loading. There are many considerations when determining torque specifications. The mass of the components, vibration and G-force requirements, and the size of the screws are just a few. When possible, the same torque should be used during the DPT operation that is specified on the assembly drawing. One of the best methods for reducing the amount of torsional stresses created during the torquing procedure is to use a 3 click - back off - 1 click torque sequence. The back off step relieves residual stresses on the threads that could eventually relieve themselves during the machining operation and cause surface figure errors.

FUTURE DIRECTION

There are many opportunities for continued improvement in reliability, cost reduction and improved performance of diamond point machined snap-together optics. As new castable materials are developed, the future of snap-together optics appears to be headed in the direction of fewer and fewer components. Mirrors will be designed as an extension of gimbles and detector mounts, and the number of system components will decrease as reliability and system performance increase. The demand for lightweight airborne systems will continue to create a need to develop more effective cutting methods, innovative stress-free mounting techniques, and alternative substrate materials such as lightweight metal-matrix composites. Diamond point machining is still a relatively new technology and current system designs are just now beginning to realize the benefits that diamond point machining offers. Diamond point machined snap-together optics have proven to be producible in production quantities and infrared system designs which fully utilize this technology.

References:

1) Texas Instruments, Optics Manufacturing and Materials, "Optics Design Guide", Section VII, 1985.
2) John Schaefer and Steve Scott, "Single-Point Diamond Machining of Optical Components", Texas Instruments Technical Journal, February 1991.
3) Daniel Vukobratovich, "Introduction to Opto-Mechanical Design", Optical Sciences Center, University of Arizona, 1986.

The Effects of Fundamental Axial Spindle Motion on the Optical Performance of Diamond Turned Surfaces

Mark Craig Gerchman

Rank Taylor Hobson Inc.
Keene, New Hampshire 03431 USA

Presentation Abstract†

Recent advances in machine tools for single point diamond turning have significantly improved the quality of machined surfaces. The most recent generation of SPDT machines are producing optical surfaces suitable for many visible wavelength applications. It is beneficial to analyze these surfaces to isolate the sources of residual errors with the view to further improvement.

This presentation will examine the surface errors which result as a consequence of fundamental axial spindle motion in the diamond machining process. In addition, the effect of this error on the optical performance of these surfaces will be investigated. A mathematical model of this error will be developed and shown to be in good agreement with experiment results.

Fundamental axial spindle motion consist of a once-per-revolution axial movement of the work holding spindle. This motion typically results from alignment errors in the spindle's thrust bearing geometry. Although this motion is small (typically less than 50 nm for air bearing SPDT lathes) it creates a unique signature error on the surface of the diamond turned component. Because the error is inherent in the machine tool's geometry, it's repeatability permits modeling so the effect on the optical performance of SPDT components can be investigated.

An analytical model for the error created by fundamental axial spindle motion is simple. In a cylindrical coordinate system aligned with the spindle motion the sagitta surface error can be expressed as: $z = b \sin(\varphi)$; where b is the amplitude of the fundamental spindle axial motion and φ is the azimuthal angle. Figure 1 shows this surface error referenced to a surrounding plane. The pole discontinuity at the vertex of the surface has an important consequence. The lack of a well behaved paraxial region makes traditional image analysis difficult. In addition, ray-tracing from a surface with a singularity poses unusual challenges for most lens design tolerancing programs.

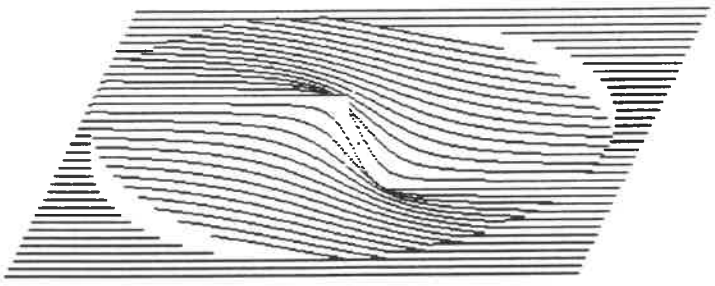


Figure 1
Surface Error Due To Fundamental Axial Motion

The fitting of optical fabrication errors is most frequently accomplished by the application of Zernike circle polynomials. In addition to orthogonality and invariance over the unit circle these polynomials are well suited to the evaluation of optical systems because their numerical similarity with the balancing of optical aberrations of different orders. The use of Zernike circle polynomials is also common in ray-tracing optical design software. Although the Zernike's have favorable properties for the representation of wavefront and surface errors, their ability to represent a pathological surface (e.g. a surface with fundamental axial spindle motion) with a restricted number of terms will be examined.

Based on a symmetry argument, the least-squares fitting of the lowest 36 Zernike terms to this simple model involves the use of terms with only fundamental sinusoidal dependence. The five Zernike polynomial terms that have a sin(φ) dependence are shown to the right. Because of the pole singularity, the integral least-squares matrix for these five Zernike coefficients is ill-conditioned. However, a transformation to another set of polynomials permits an exact solution. The values for the Zernike coefficients are given above and their similarity to Fourier terms for step functions is not coincidental.

$$
\begin{aligned}
z = {} & a_2\,\rho\sin(\varphi) \\
& + a_7\,\rho\sin(\varphi)\,(3\rho^2-2) \\
& + a_{14}\rho\sin(\varphi)\,(10\rho^4-12\rho^2+3) \\
& + a_{23}\rho\sin(\varphi)\,(35\rho^6-60\rho^4+30\rho^2-4) \\
& + a_{34}\rho\sin(\varphi)\,(126\rho^8-280\rho^6+210\rho^4-60\rho^2+5)
\end{aligned}
$$

$$
\begin{aligned}
a_2 &= 1.333333 = (1*4)/(1*3) \\
a_7 &= -0.533333 = (-2*4)/(3*5) \\
a_{14} &= 0.342857 = (3*4)/(5*7) \\
a_{23} &= -0.253968 = (-4*4)/(7*9) \\
a_{34} &= 0.202020 = (5*4)/(9*11)
\end{aligned}
$$

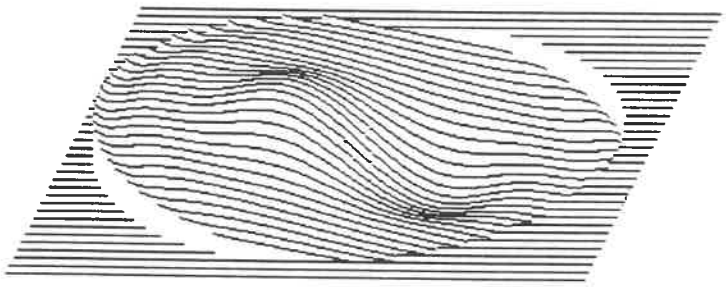


Figure 2
Surface Errors Represented By Zernike Fit

Figure 2 shows the Zernike circle polynomial fit using these coefficients. Although the overall form is similar, there are clearly an insufficient number of terms to fully represent the error. In particular, the slope within the paraxial region can not, with any reasonable number of Zernike terms, represent the effects of the pole singularity.

This presentation will conclude by comparing the exact representation of error due to fundamental axial spindle motion to the Zernike circle polynomial fit with regard to optical performance. This comparison will include both an analysis by exact ray-tracing techniques and by computation of the Strehl intensity for these functional relationships.

---

†The full manuscript for this presentation is to be published as "Fundamental Axial Spindle Motions and Optical Tolerancing", Proc. SPIE Vol. 1573 - Commercial Applications of Precision Manufacturing at the Sub-micron Level (SIRA - London), 1991.

DEVELOPMENT OF DRIVE-SYSTEMS FOR HIGH ACCURACY ROTARY POSITIONING DEVICES

M. Weck, J. Luderich
Fraunhofer-Institut für Produktionstechnologie
IPT, Aachen, Germany

Rotary tables, circular dividing systems and other high precision angle generators require drive systems with the ability to position the axis of rotation over a range of 360 degree with a repeatability in the order of 0.01 arc seconds. Typical specifications for these drive systems are the repeatable positioning to ± 1 digit, high ratio of minimum to maximum rotational speed and a non-influencing coupling of the drive system 1to the rotary device. Such systems have to be able to position the axis of rotation with a repeatability of the resolution of the rotatry encoder.

There a numbers of alternative actuator concepts that may be used for high performance drive systems (Fig 1). These actuator concepts can be devided principly in two main groups: systems with one single and systems with two actuators. Other criterions are the use of a gear and the position of the drive system relativly to the rotary table.

In this report three general concepts and their qualification for high accuracy rotary positioning devices are presented and discussed.

- direct drive and servo system
- friction gear drive
- piezo actuator for small angles in combination with another drive for large rotary displacements

Fig. 2 shows the concept of the direct drive and servo system. A high resolution DC torque motor is used

as actuator. The rotating speed is measured by a DC Tachgenerator, the position by a high resolution rotary encoder. The most accurate commercial available rotary encoder with 36.000 lines per revolution in combination with an interpolation unit, which multiplies the signal of the rotary encoder by a factor of 1024, is used. This leads to a resolution of 0,035 arcsec and an accuracy of 0,2 arcsec.

The direct drive works under closed loop control. While the current control has been realised analog in the servo-amplifier, the velocity and the position control works digital in a computer. Additional to the velocity and position control, control functions and compensation-algorithms are realised in the computer. The use of compensation algorithm gives the possibility to correct e.g. thermal drifts within in the servoamplifier or to compensate long wave angular errors of the rotary encoder to get a higher accuracy. Fig.3 shows the result of an typical positioning procedure.

Together with the direct drive system the qualification of a friction gear drive with a reduction of i=100 and a rotatry positioning device with a piezo translator has been examined.

Main advantage of the friction drive is it's use as a reduction gear, which allows the employment of simple and cheap motor drive systems. Piezo translators are commercialy available components and offers the possibility of highest resolution. Both positioning devices require a relative complex mechanical design which can lead to dynamic problems.

In Fig. 4 shows the mechanical concept of the test set up for the friction gear drive and the integration of the piezo translator on principle.

The geometrical and positioning behaviour of these three systems has been measured. Measurement results of the positioning behaviour of the different drive systems will be presented in the oral report.

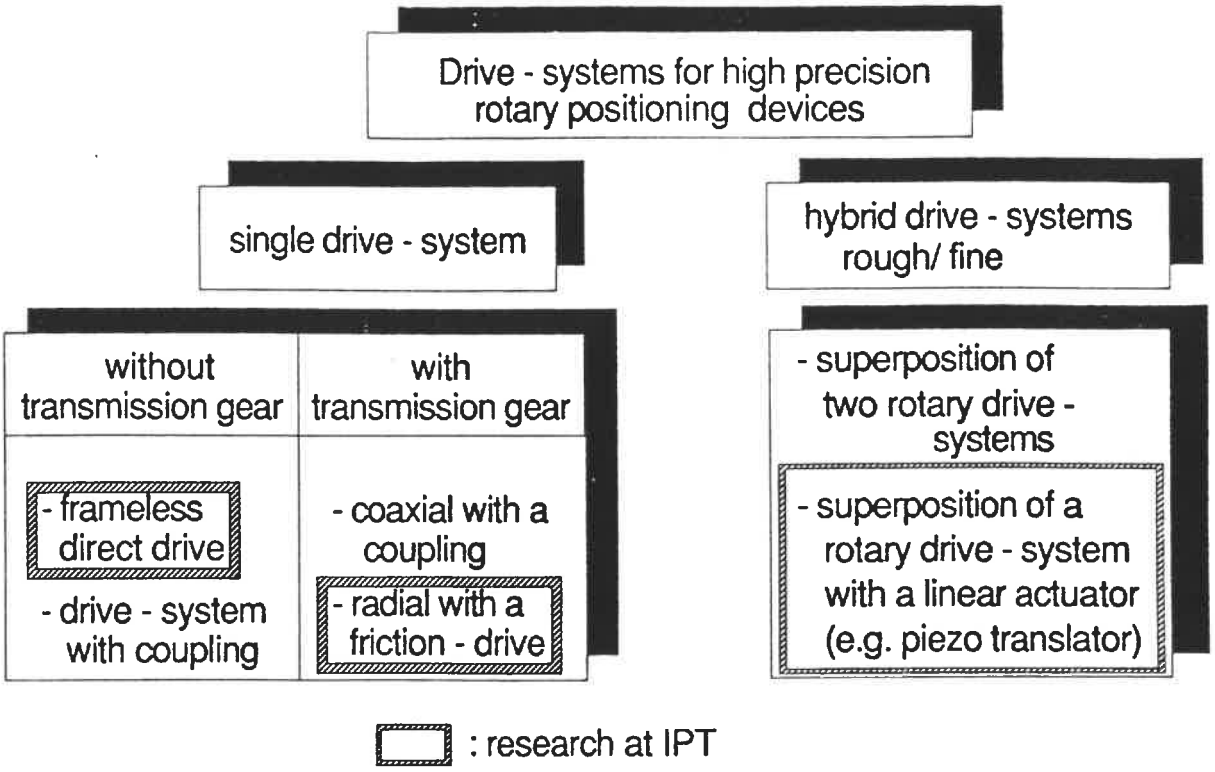

Drive - systems for high precision rotary positioning devices
single drive - system
hybrid drive - systems rough/ fine
without transmission gear
with transmission gear
- frameless direct drive
- drive - system with coupling
- coaxial with a coupling
- radial with a friction - drive
- superposition of two rotary drive - systems
- superposition of a rotary drive - system with a linear actuator (e.g. piezo translator)
: research at IPT

Figure 1: Drive systems for high precision rotary positioning devices

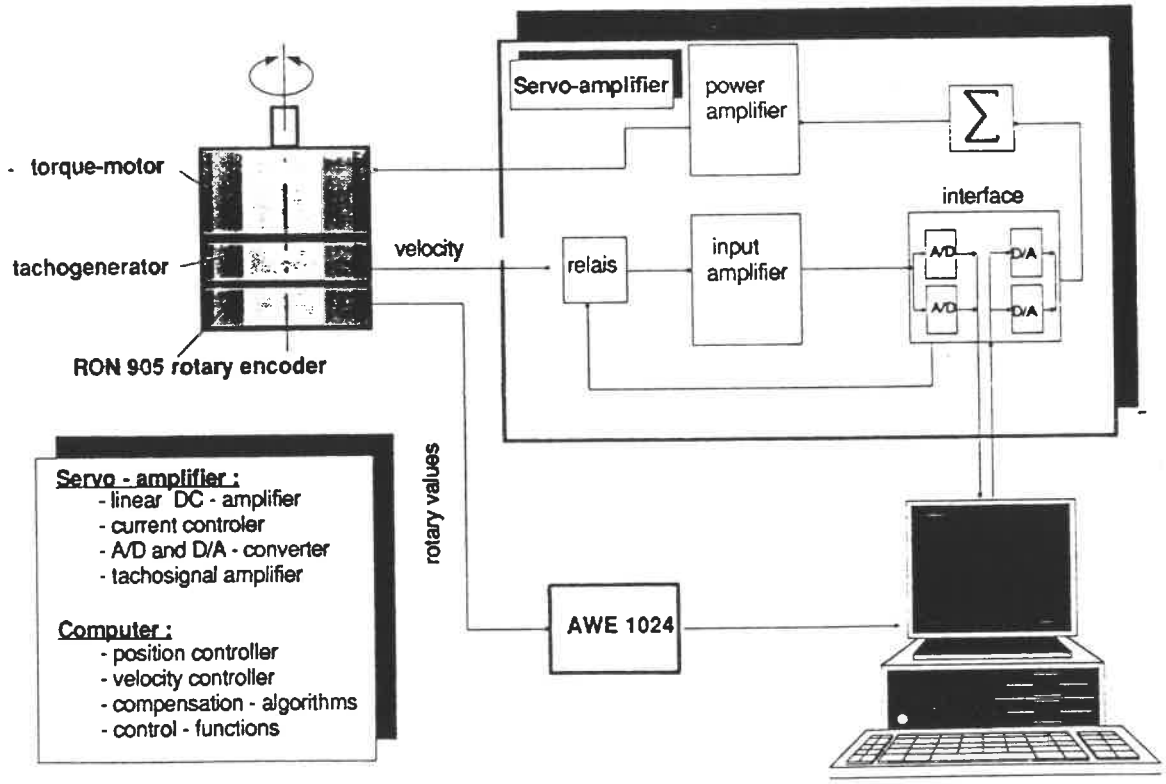


Figure 2: Concept of the direct drive and servo system

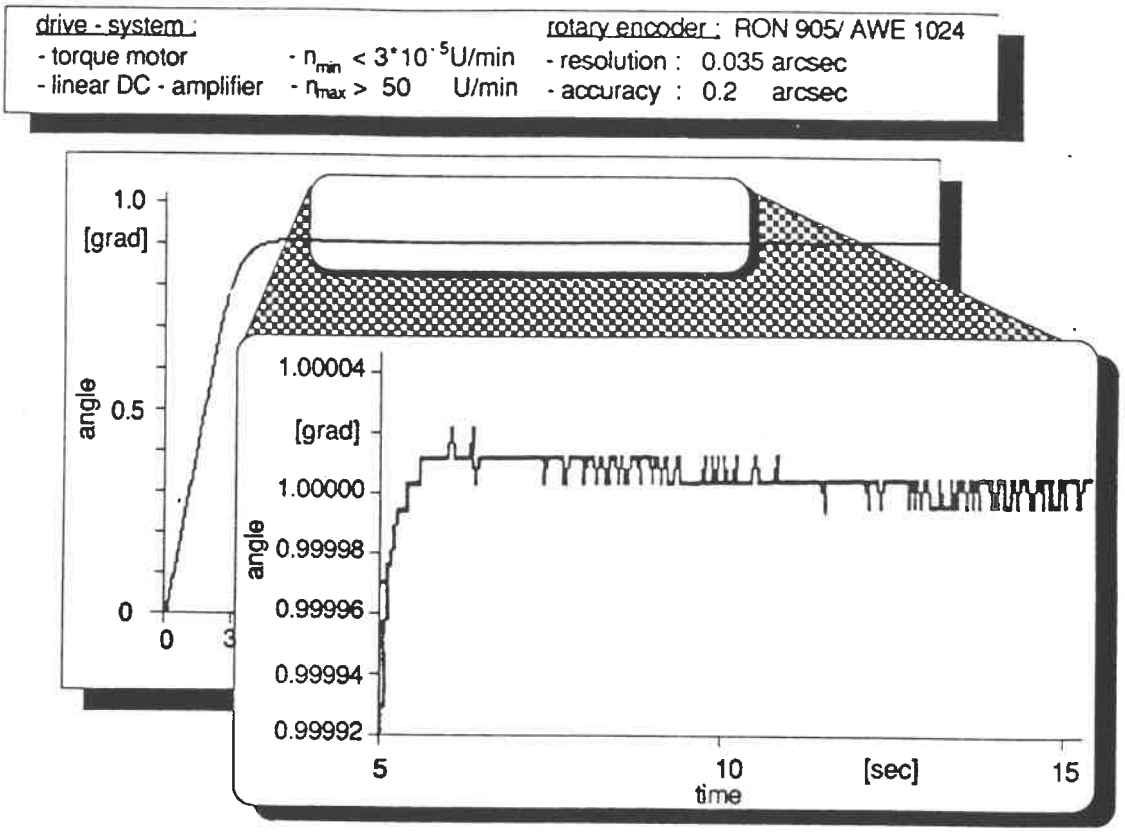


Figure 3: Measurement result of an typical positioning procedure of the direct drive system

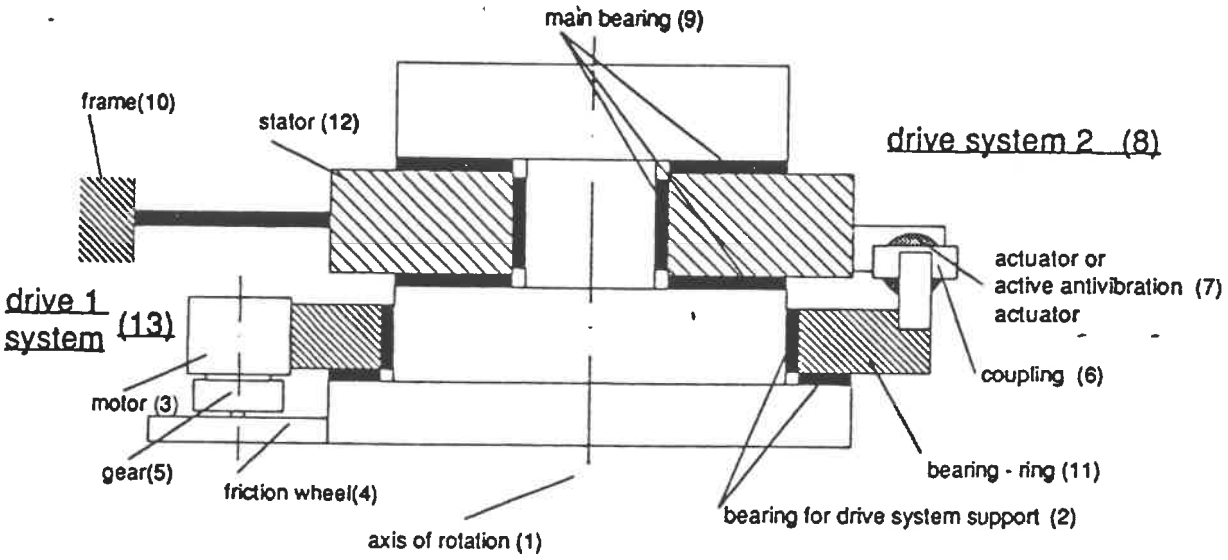


Figure 4: Mechanical concept of the test set up of the friction gear drive and the integration of the piezo translator on principle

Feed and Frictional Load Characteristics of Carriage on Plain Bearing Guideways in the Range of Nanometer Positioning

1991 Annual Conference of American Society for Precision Engineering
Santa Fe, New Mexico, USA, October 13-18, 1991

Akira Kanai*, Junichi Yoshioka**, Masakazu Miyashita*
*Ashikaga Institute of Technology, **NUTR K.K., Japan

1. INTRODUCTION

Ultraprecision and high stiffness grinding machine for ductile grinding of brittle materials has to be provided with feed mechanism with a high resolution of order of nm and a stiffness of order of N/nm in feed direction.

In this paper, a testing of the newly proposed high stiffness force operated linear actuator for a carriage on plain bearing guideways is applied to carrying a series of experiments on the feed characteristics and frictional load characteristics of 150 kg carriage approximately in the travelrange of micrometer to nanometer.

The experimental results say that a feed resolution of the carriage under several 100 N nominal frictional load is one nanometer and the frictional load is composed of the semi-linear spring load domain and the semi-stationary frictional load domain. And the very fine stick-slip phenomena of approximately 100 nm/step are observed experimentally according to the selected design parameters.

2. TEST RIG AND EXPERIMENTAL CONDITIONS

Figure 1 shows the construction of the test rig. The carriage is hydrostatically supported in vertical direction and is supported by the composite bearing guideways, which are couples of hydrostatic pads and plain bearings, in horizontal direction. Here, the hydrostatic pads function to control the plain bearing loads and realize the uniform contact pressure distribution of the bearings along the guideways. The main experimental conditions are as follows :

Mass of carriage : 150 kg approx.
Total area of plain bearing : 295 cm^2
Total equivalent area of hydrostatic pads : 119 cm^2
Supply pressure to hydrostatic pads, P_h : (1) 392 kPa, (2) 294 kPa, (3) 196 kPa
Total load on plain bearings : (1) 2.16 kN, (2) 1.66 kN, (3) 1.2 kN - estimated from the measured pocket pressure of a hydrostatic pad
Plain bearing load, P_s : (1) 73 kPa, (2) 57 kPa, (3) 41 kPa
Semi-stationary frictional load : (1) 320 N, (2) 270 N, (3) 145 N
Sensitivity of feedback sensor - capacitive displacement detector : 400 mV/um

Figure 2 is the photo of the force operated linear actuator. Figure 3 is the photo of the test rig equipped with the actuator. The piston of the actuator is hydrostatically supported and no seal is applied, and the pressure difference between the cavities can be the applied load on the actuator. The major performances of the positioning feedback system involving the force operated linear actuator without external load are as follows :

Feed resolution : 0.05 nm, which means the 20 feed steps of 0.05 nm/step are equal to 1 nm.
Maximum static stiffness(measured) : 66 N/nm
Band width : 100 Hz

3. EXPERIMENTAL RESULTS

3.1 Feed characteristics for stepwise feed - feed resolution

Figure 4 shows the feed characteristics of the carriage under the plain bearing load of 73 kPa. They are the characteristics of (a) 100 nm/step, (b) 10 nm/step and (c) 1 nm/step.

Figure 5 is the feed characteristics for 1 nm/step feed under the bearing load of 41 kPa. These test results show that the feed resolution of the carriage is 1 nm or the less. In viewpoint of S/N ratio, the bearing load of 41 kPa can be considered to give a higher feed resolution.

3.2 Feed characteristics for constant feed rate and frictional load characteristics of plain bearing

The positioning characteristics and the frictional load characteristics of plain bearing for a reciprocating movement of the carriage at the constant feed rate of 0.16 um/s with a period of 25 seconds are shown as follows :

(1) In case of the plain bearing load P_s = 73 kPa,

Figure 6(a) and (b) are records of the positioning of carriage and the frictional load of the bearing.

The experimental conditions are :

Total load on the bearings : 2.16 kN

Averaged semi-stationary frictional load for reciprocating movement : 320 N

The positioning characteristics :

no lost motion at the turning point in the reciprocating movement

The frictional load characteristics :

Figure 6(b) shows that the frictional load is composed of the semi-linear spring load domain and the semi-stationary frictional load domain. The former is a transient one.

The averaged equivalent linear spring stiffness for the displacement of 100 nm from the turning point is 5 N/nm approx..

(2) In the case of the plain bearing load P_s = 41 kPa,

Figure 7(a) and (b) are records of the positioning of the carriage and the frictional load of the bearing.

The experimental conditions are :

Total load on the bearings : 1.2 kN

Averaged semi-stationary frictional load for reciprocating movement : 145 N

The positioning characteristics : no lost motion

The frictional load characteristics :

The averaged equivalent linear spring stiffness for displacement of 100 nm from the turning point is 2.3 N/nm approx..

3.3 Positioning and frictional load characteristics for stepwise feed

Figure 8(a), (b), and (c) are records of the positioning characteristics and the frictional load characteristics for stepwise feed. They are the characteristics for (a) 50 nm/step, (b) 10 nm/step and (c) 1 nm/step.

The experimental conditions are :

Supply pressure to the pads P_h = 294 kPa

Plain bearing load P_s = 57 kPa

Total load on the plain bearings : 1.66 kN

From the experimental results it is shown that the frictional load characteristics is composed of the following two domains :

(1) the domain of equivalent semi-linear spring characteristics, and

(2) the domain of semi-stationary frictional load characteristics.

The ratio of the increased part of frictional load to the stepwise displacement is defined an equivalent linear spring stiffness. The experimentally derived relationship between the accumulated stepwise displacement and the equivalent linear spring stiffness is shown in Figure 9. A tendency of lower equivalent linear spring stiffness for larger stepwise feed is due to a soft spring characteristics as a whole of the domain of equivalent linear spring load characteristics.

3.4 Observed stick-slip phenomena at a constant feed rate of 0.16 um/s

In a series of experiments, the very fine stick-slip phenomena are recorded in the composite bearing guideway system according to selected design parameters of the system. Figure 10 is an example.

In the range of about 230 nm displacement from the turning point no stick-slip occurs and then the stick-slips grow up to the maximum load variation of 104 N. The displacement of one slip is almost constant of 76 nm approx. and is repeated regularly similar to a stepwise feed. Such generation of the very fine stick-slip phenomena is dependent of the design parameters such as the plain bearing load, a loop gain of feedback control system, a location of the feedback sensor on the carriage attached and so on.

4. CONCLUSION

(1) A feed resolution of the 150 kg carriage guided by the plain bearing under nominal frictional load of 320 N is 1 nm or the less.

(2) Frictional load characteristics is composed of two domains - semi-linear spring characteristics and semi-stationary characteristics.

(3) Very fine stick-slip phenomena of approx. 100 nm/step are experienced.

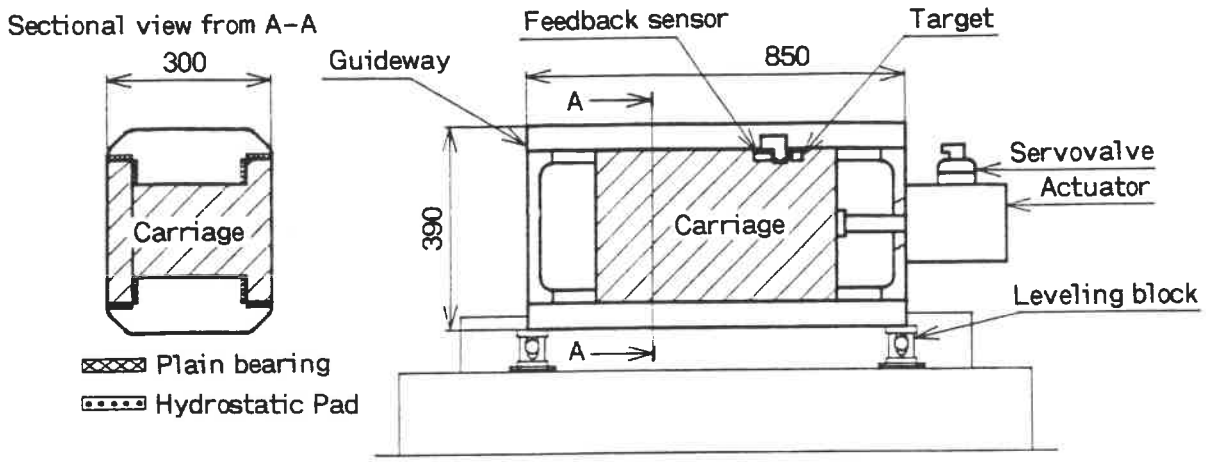


Fig.1 Construction of the test rig

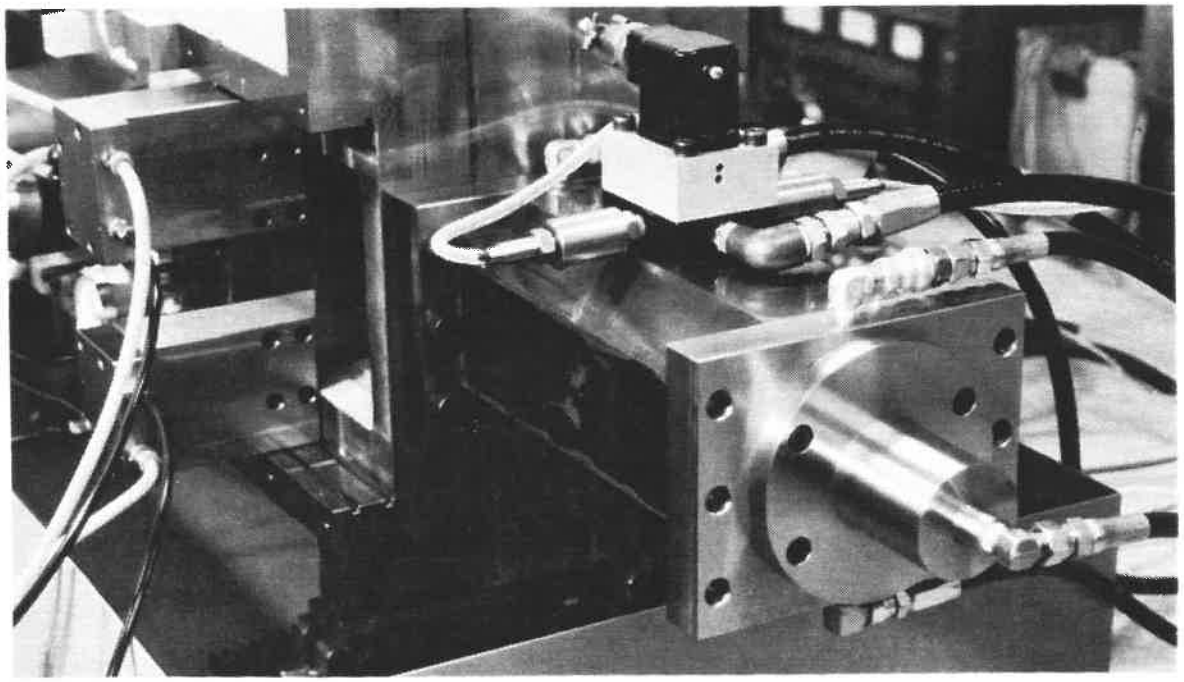


Fig.2 Force operated linear actuator

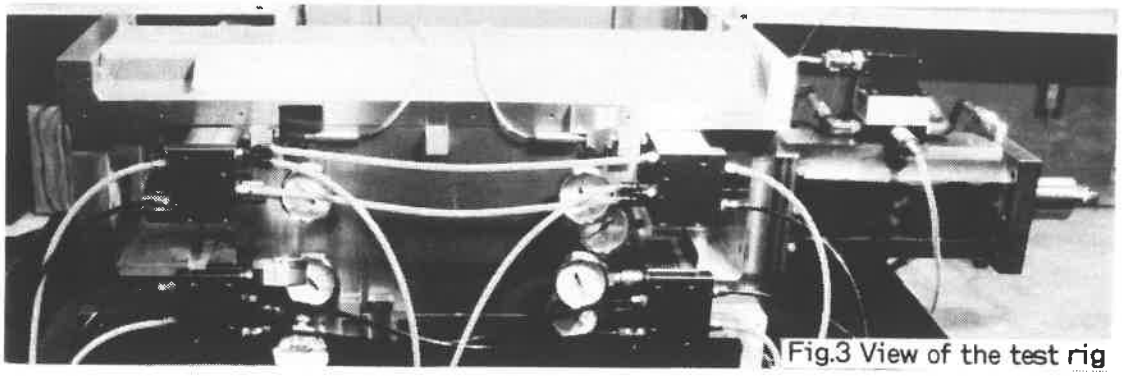


Fig.3 View of the test rig

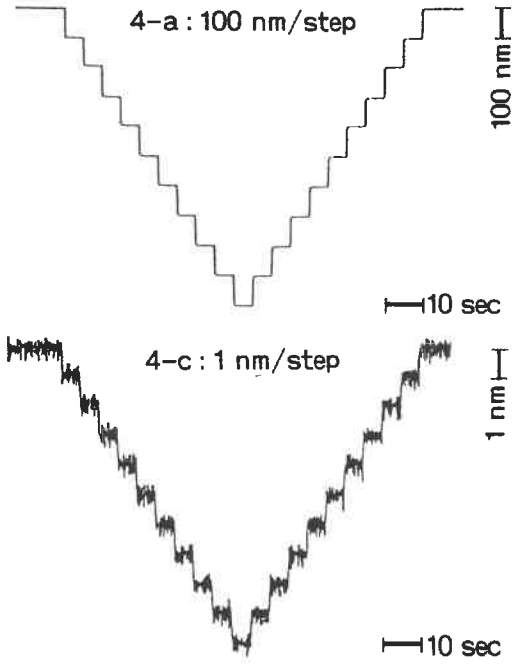


Fig.4 Feed resolution (Ps=73 kPa)

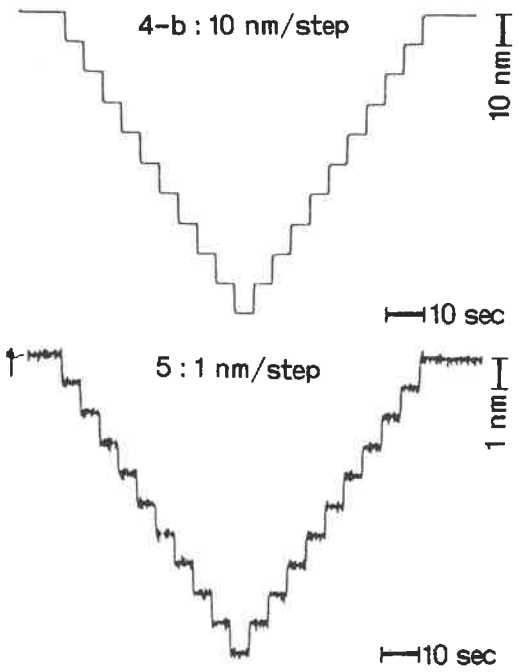


Fig.5 Feed resolution (Ps=41 kPa)

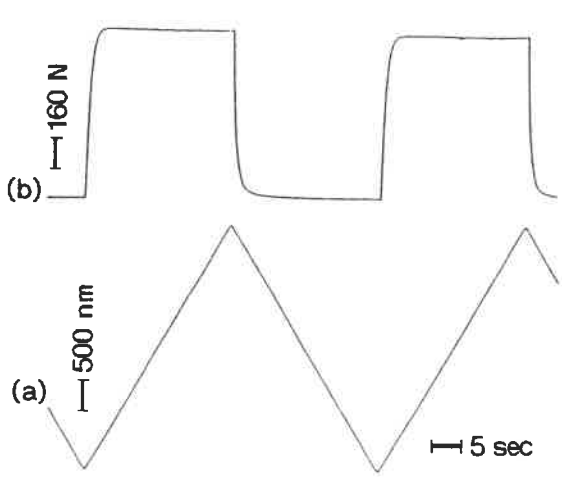


Fig.6 Displacement of carriage (a) and frictional load (b) for the constant feed rate (Ph=73 kPa)

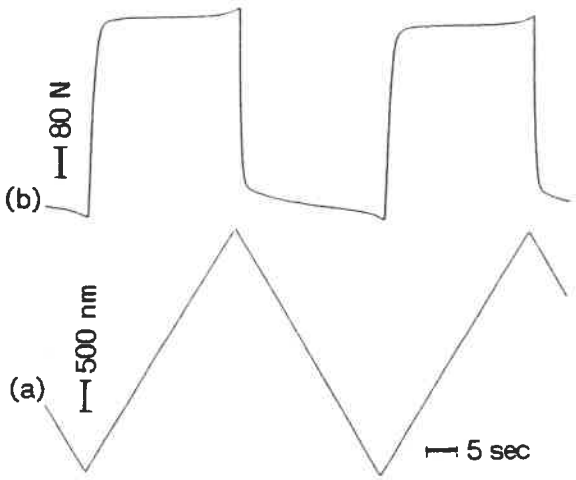


Fig.7 Displacement of carriage (a) and frictional load (b) for the constant feed rate (Ph=41 kPa)

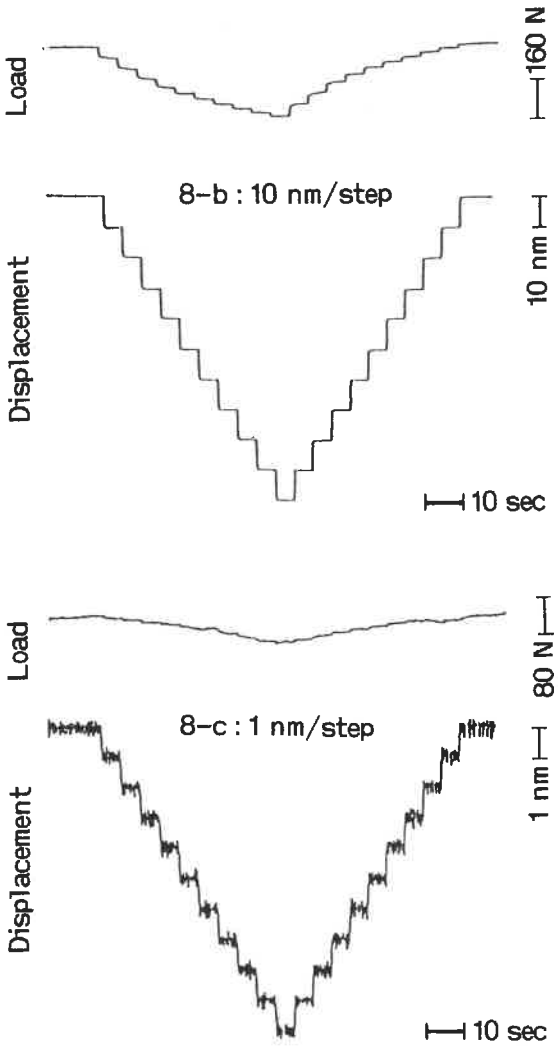


Fig.8 Frictional load for stepwise feed (Ph=57 kPa)

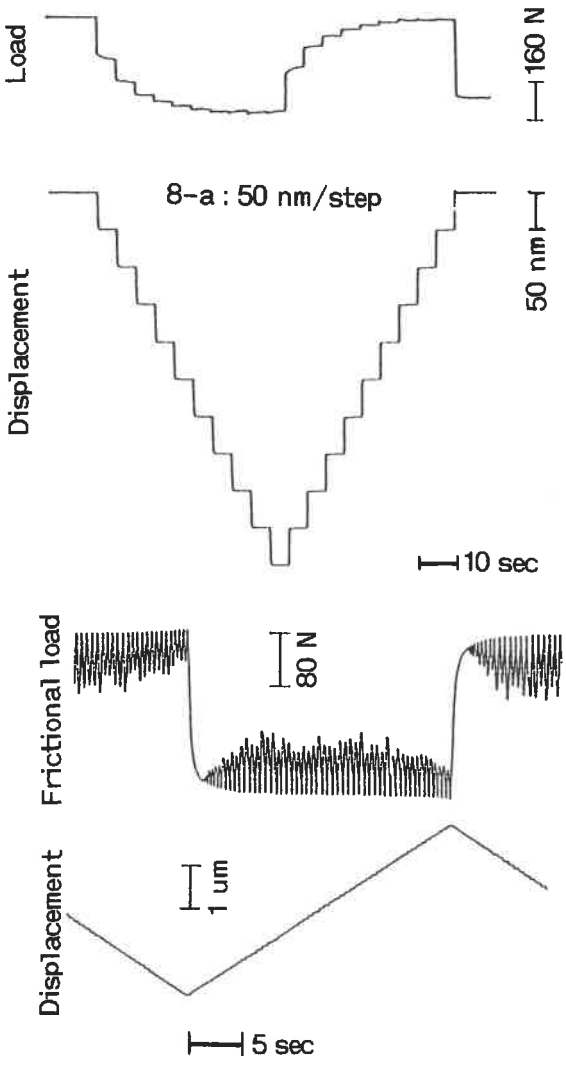


Fig.10 Example of stick-slip phenomena

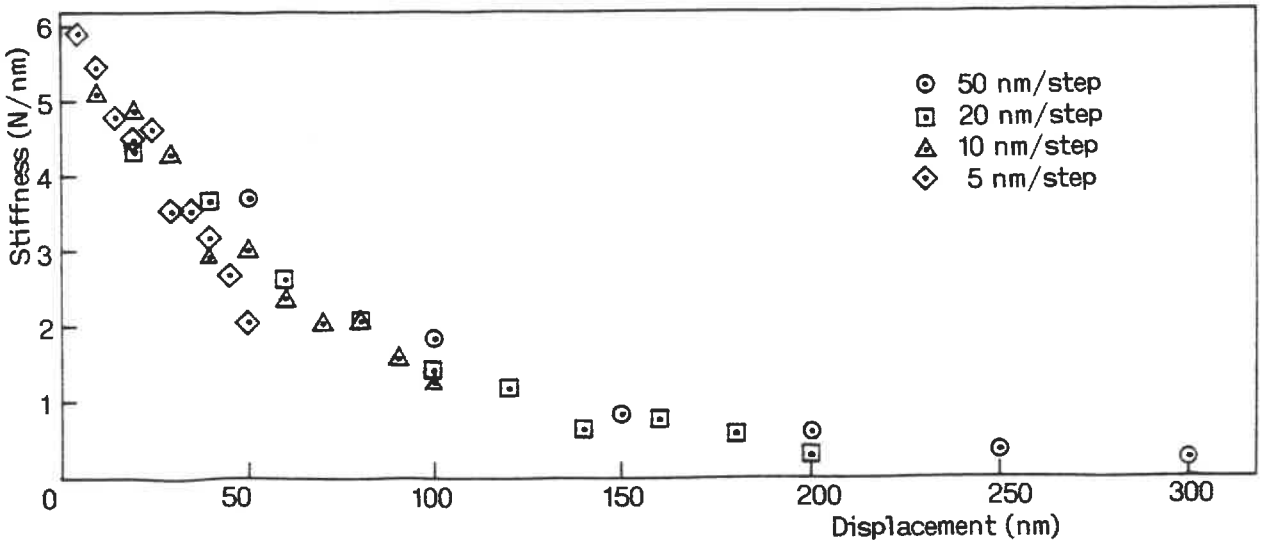


Fig.9 Equivalent spring stiffness to accumulated stepwise displacement (Ph=57 kPa)

ULTRA-PRECISE AC LINEAR MOTOR FOR NANOMETER POSITION CONTROL

SHIGERU FUTAMI and AKIHIRO FURUTANI

YASKAWA Electric Corp. Tsukuba Research Laboratory
5-9-4 Tohkohdai, Tsukuba, 300-26 Japan

1. Introduction

Studies on nanometer position control are being vigorously pursued. This has recently lead to the development of a few practical nanometer position control systems. Various drive methods such as those using friction drives [1], ball screws [2] and hydraulic actuator [3] are used in these systems.

In this paper an AC linear motor for a direct drive method providing nanometer position control is described. The direct drive method has several advantages compared to the other methods. These advantages include no contact, a simple mechanical structure, high rigidity, and quick response. The design concept for the motor that provides nanometer controllability, the actual design, and experimental testing results are described.

2. Design concept

In ultra-precise position control, especially that at the nanometer level, the selection of a drive method is extremely important. The key features of a good drive method are high force resolution, force generation in only the drive direction, high rigidity, no mechanical coupling, simple construction and small size.

Our drive method uses direct drive by an AC linear motor. This drive method has the advantages of no-contact force generation and simple mechanical structure. The features of our design concept and our choices are the following:

***No-contact force generation** (AC type linear motor)
***Minimum force ripple** (coreless-type motor and optimum electro-magnetic design based on FEM analysis)
***Maximum response** (coreless-type motor that has a minimum electric time constant)
***Force resolution** (force resolution 'df' must be less than K*dx where K equals spring constant and dx equals position control resolution. If K=10 N/micro-meter and dx=1 nm then df<10 mN)
***Minimum heat generation** (moving magnet-type motor with no heat generation in magnets because they are on the moving part of motor, and easy radiation of the heat generated in the coil by

heat transfer through the stage base on which the coil is fixed)
*Small size; (using powerful SmCo rare earth magnets)

3. Design and experiments

A coreless moving magnet-type AC linear motor was designed to according to the design concepts described in Section 2. Figure 1 shows a schematic diagram of the designed motor. Six magnet pairs are placed on the moving part of the motor. The pole pitch is 25 mm and the winding pitch is 50 mm. Samarium-Cobalt magnets are used. The size of the magnets is 9x20x28 mm. A three-phase motor winding is placed at the center of these magnet pairs as shown in the cross sectional view. The winding is solidified by resin. No magnetic core is used. A conducting tape was wounded around a non-magnetic plate in a triangular pattern. The configuration of the magnets and the winding was designed to minimize force ripples.

The results from electromagnetic FEM analysis of the motor are shown in Figure 2. The maximum force is 20 N and the force ripple (fifth order harmonics) is 0.012 % over the maximum force. No cogging force is generated because of the coreless-type motor. Therefore, the force irregularity should be negligibly small.

The AC linear motor was controlled so that the magnetic flux vector and the current vector are perpendicular to each other. The star-connection three-phase winding is driven by two current amplifiers. The two currents are controlled so that the phase difference of the electric angles is 120 degrees.

The resistance 'R' of each winding was 20.8 ohm and the inductance 'L' was 3.24 mH. The time constant of the motor 'L/R' was 0.155 ms.

Experiments showed that the cutoff frequency of the current control of the motor was over 10 kHz and the driving force was linearly generated over a 5 kHz bandwidth.

The force resolution was measured by applying a sinusoidal reference force to the spring and measuring the resulting deformation by using a sensor connected to the motor as shown in Figure 3. The spring constant was 8.5 N/micro-meter and a sinusoidal deformation less than 1 nm was generated. The force resolution was thus concluded to be less than 8.5 mN and the relative resolution was 4.25E-4 (8.5 mN/20 N). No dead-zone was observed in the sinusoidal deformation response.

Nanometer position control experiments were carried out using a one-axis stage mechanism driven by the AC linear motor [4]. Figure 4 shows the one nanometer step response of the stage. In this response the position of the stage completely followed the reference position and no backlash was observed. This experiment clearly shows that our designed motor has controllability at the nanometer level.

4. Conclusions

An AC linear motor for nanometer position control was designed and evaluated. Moving magnet, coreless-type motor was chosen to satisfy the design concepts aimed at achieving ultra-precise position control at nanometer level. Experimental results showed that the force resolution was less than 10 mN over 5 kHz bandwidth. Position control experiments showed that the motor has position controllability of less than 1 nm.

Acknowledgement

A part of this study was carried out in the Yoshida Nano-Mechanism project in the ERATO program of the Research Development Corporation of Japan.

We would like to thank Mr. Hiroshi Ikeda and Mr. Atushi Kawahara for their original design of the Gap Winding Motor used in this study.

References

[1] P. B. Leadbeater: A unique machine for grinding large, off-axis optical component, PRECISION ENGINEERING, Vol.11 No.4 (1989)
[2] Y. Kawase: Ultra-precision Positioning Using Leadscrew Drive (2nd Report), Journal of the JSPE, Vol.57 No.7 (1991)
[3] A. Kanai: Ultra-precise position control using a force-controlled hydraulic actuator, 1986 Autumn Conf. of JSPE (1986)
[4] S. Futami: Nanometer positioning and its micro-dynamics, Nanotechnology, Vol.1 No.1 (1990)

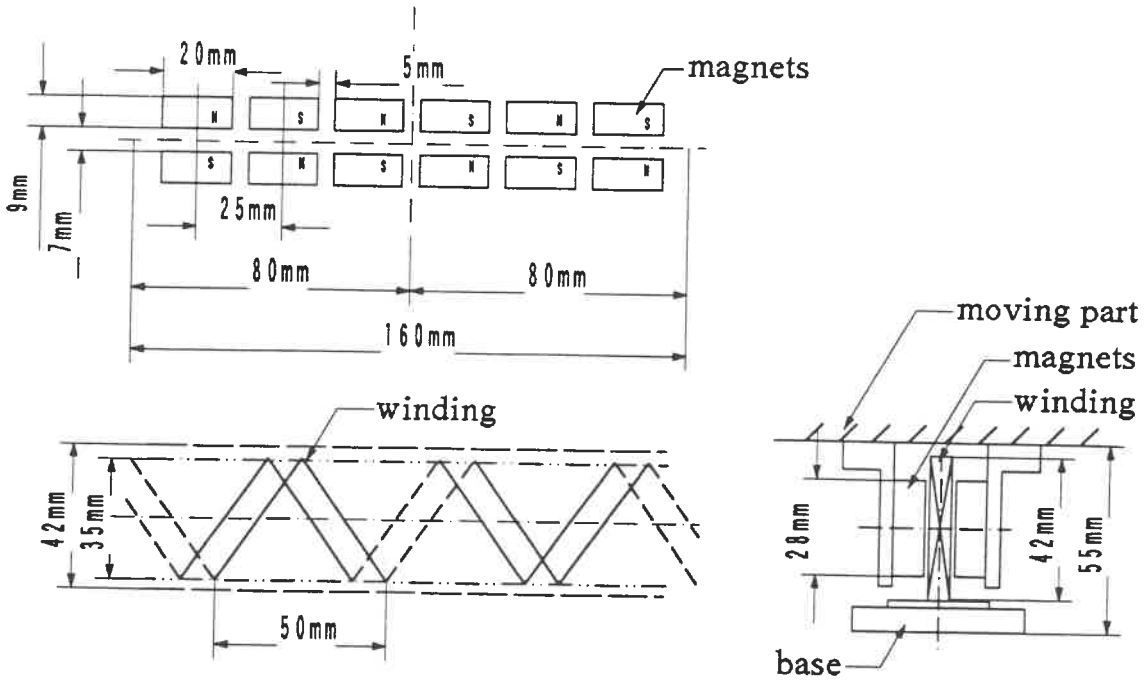


Figure 1 Schematic diagram of designed AC linear motor.

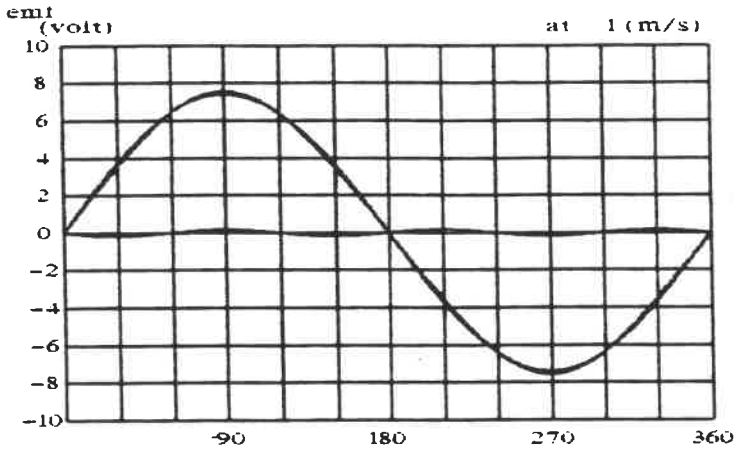


Figure 2 Force and force ripple results obtained from electromagnetic FEM analysis.

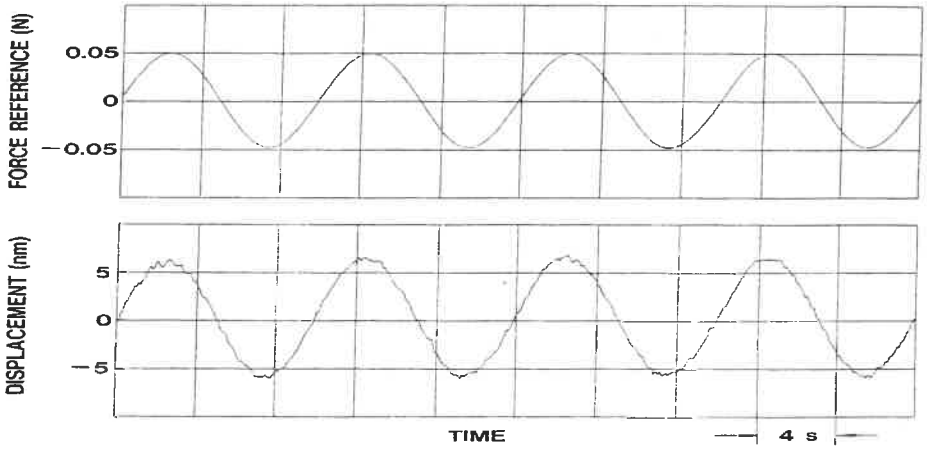


Figure 3 Measurement of force resolution of the motor by applying reference force to a spring (10 N/micro-meter) and measuring the resulting deformation.

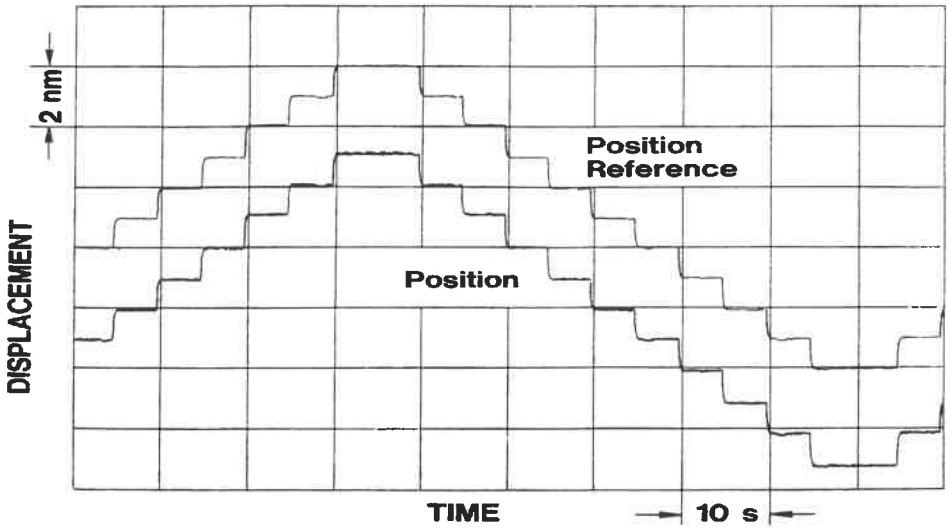


Figure 4 One nanometer step response of a one-axis stage driven by the designed AC linear motor.

MODEL REFERENCE ADAPTIVE CONTROL OF NONLINEAR MICRODYNAMICS OF A BALL-SCREW DRIVEN PRECISION SLIDE

Peter I. Hubbel
Paul I. Ro
Precision Engineering Center
North Carolina State University
Raleigh, NC 27695-7918

Introduction

The increasing demands that precision engineering applications place on positioning systems has prompted an investigation into the nanometer positioning capabilities of ball-screws. Though ball-screws have been used for many years, they traditionally were not applied to problems requiring sub-micrometer accuracy.

Standard feedback control schemes such as PID can improve positioning accuracy, but the dynamics of the ball-screw at the sub-micrometer level are nonlinear and advanced control design techniques are necessary to obtain further improvements in accuracy.

This paper describes the application of a model reference adaptive controller (MRAC) for sub-micrometer motion. The controller design process is reviewed, including system characterization, adaptive control law derivation, and experimental results.

Characterization

Open loop testing indicates that small angle ball-screw rotation (and subsequent slide displacement) occurs for any applied torque.[1] This "spring" behavior, resulting from elastic deformation of the contact patches on the balls, allows for nanometer motion by the application of a low torque.[2] However, as the applied torque increases, a friction breakaway is observed and the deformation remains constant. The plant model subsequently includes two different modes, one which applies prior to the friction breakaway and another which applies after the breakaway.

The pre-breakaway dynamics (microdynamics) are dominated by the observed elastic behavior. Although the measurement obtained by Cuttino in Figure 1 indicates that the torque-displacement relationship is nonlinear, the model assumes linearity in order to provide for the application of well-developed control design techniques[2,3]. The elastic behavior is modelled as a deforming element which links the ball-screw to the friction interface. The post-breakaway dynamics (macrodynamics) are based on a classical dynamic friction model and assume that the deformation of the microdynamic "spring" is constant.

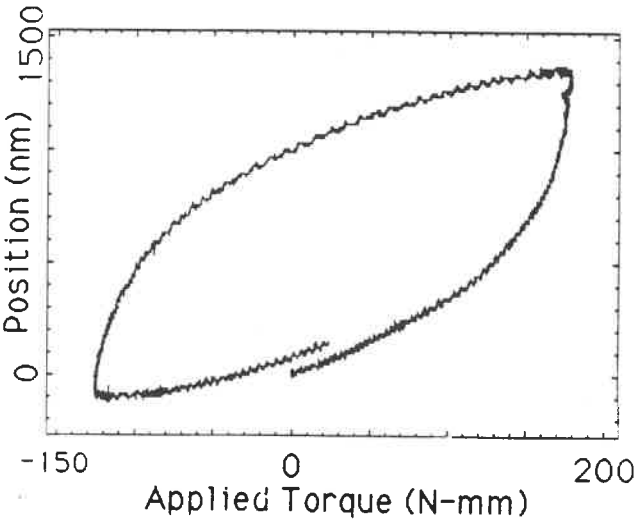


Figure 1: Slide displacement vs applied torque for small displacement.

The friction level of the air-bearing slideway is negligible compared to that of the rotational elements, so a second order rotational model can be used to illustrate the friction induced dynamics. The model is applicable to this particular apparatus when the nut-slide coupling is sufficiently rigid. Figure 2 illustrates the structure of the model, showing

that θ_{mdp} is the position at the friction interface relative to ground.

The dynamics are described by

$$J\ddot{\theta}_p+B\dot{\theta}_p+B_{fr}(\dot{\theta}_p-\dot{\theta}_{mdp})+K_{fr}(\theta_p-\theta_{mdp}) = u_l+u_a$$

regardless of the status of the friction interface. In the preceeding equation J is the rotational inertia of the motor armature and ball screw, B is the damping coefficient of the motor and ball-screw with respect to ground, B_{fr} is the damping coefficient of the ball-screw with respect to the microdynamic reference position θ_{mdp}, K_{fr} is the microdynamic spring constant, and θ_p is the position of the ball screw. The displacement of the slide is

$$x = K_p\theta_p,$$

where the pitch of the screw is K_p.

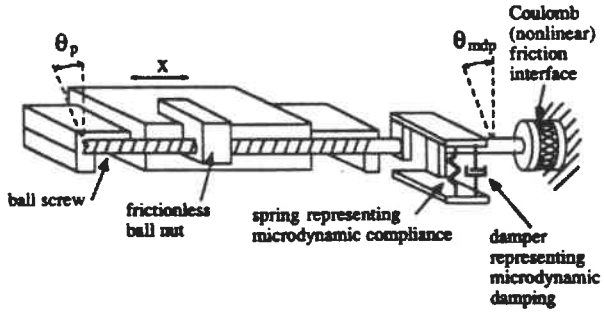


Figure 2: Conceptual diagram illustrating the structure of the model.

The specific dynamic modes are determined according to

$$\dot{\theta}_{mdp} = \begin{cases} 0 & \text{(microdynamics)} \\ \dot{\theta}_p & \text{(macrodynamics)} \end{cases}$$

When the plant is in the macrodynamic mode, the deformation of the microdynamic spring is constant and the friction torque is

$$\tau_f = K_{fr}(\theta_p - \theta_{mdp})$$

A PI control law was implemented to determine the effect the microdynamic nonlinearity has on the closed-loop response. The gains were adjusted experimentally to provide a rise time of approximately 0.1 seconds and a slightly overdamped response to a 300 nm step command. However, when the same controller gains are used for different step commands, the closed-loop response varies significantly. Figures 3 and 4 show the variations in the closed-loop response.

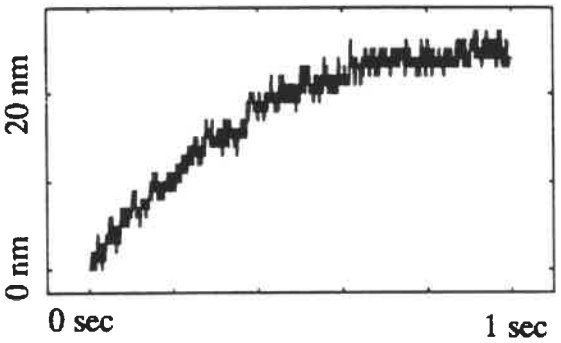


Figure 3: Response of plant to 25 nm step command while under PI control.

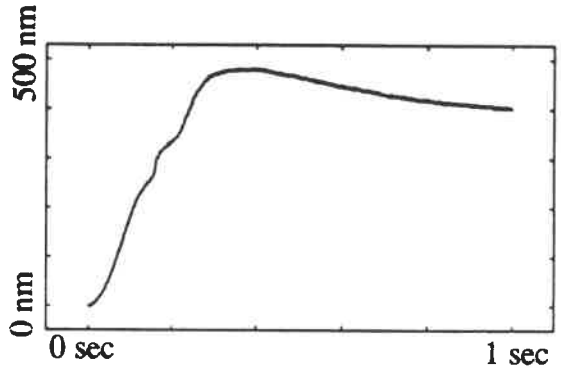


Figure 4: Response of plant to 400 nm commanded step while under PI control.

Controller Design Objective

Given the magnitude dependent dynamics exhibited in Figures 3 and 4, the objective of the control design is to provide a magnitude independent, predictable closed loop response for nanometer motion.

Model Reference Adaptive Control

The MRAC technique includes a reference model which generates an ideal response and adaptive laws which adjust variable parameters to ensure that the plant output follows the reference model output, regardless of unknown plant parameters[4]. Restrictions do apply to the

structure of the reference model and lack of plant states can hinder implementation of the adaptive laws. However, applications have shown the effectiveness of the technique.[5]

Design of Microdynamic Adaptive Controller

The microdynamic plant model includes three parameters that are critical to controller design but are not well known and may be time-varying. These are the microdynamic reference position (θ_{mdp}), the microdynamic spring constant (K_{frp}), and the microdynamic damping coefficient (B_{frp}).

The uncertainty of K_{frp} results from the linearization of the nonlinear displacement-torque relationship in Figure 1. The microdynamic damping coefficient is difficult to measure experimentally, and may also be magnitude dependent. Likewise, the reference position (θ_{mdp}) is also difficult to measure.

The MRAC scheme is designed to force the plant to follow the reference model despite poor knowledge of these parameters. Figure 5 is a block diagram which illustrates the structure of the controller, where the linear controller could be a state variable controller as all states of the reference model are observable.

The value of K_{frm} was obtained by linearization of the torque-displacement relationship and the initial estimate of B_{frm} was obtained by matching experimental closed-loop response data to computer simulation output. The reference position (θ_{mdm}) can be approximated by the ball-screw position (θ_p) when the applied torque (u_l+u_a) is zero.

The dynamics of the amplifier, armature circuit, and current regulator are neglected, as they are much faster than the mechanical dynamics of the plant. This assumption allows the output of the linear controller (u_l) and the output of the adaptive scheme (u_a) to be modelled as torque inputs to the reference model and plant.

The reference model is

$$J\ddot{\theta}_m + (B + B_{frm})\dot{\theta}_m + K_{frm}(\theta_m - \theta_{mdm}) = u_l$$

while the microdynamic plant is

$$J\ddot{\theta}_p + (B + B_{frp})\dot{\theta}_p + K_{frp}(\theta_p - \theta_{mdp}) = u_l + u_a$$

The reference model-following error

$$e(t) = \theta_m(t) - \theta_p(t)$$

is forced to go to zero by using a Lyapunov design procedure to generate the update equations for three estimated parameters which correspond to the differences between the model parameters (K_{frm}, B_{frm}, θ_{mdm}) and the unknown plant parameters (K_{frp}, B_{frp}, θ_{mdp}).[6] The adaptive input (u_a) is then defined in terms of these three estimated parameters.

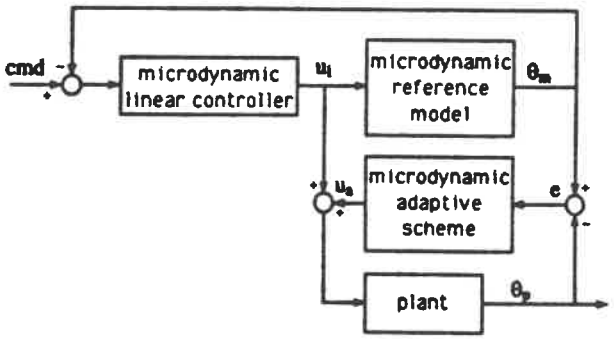


Figure 5: Block diagram of the microdynamic MRAC scheme.

Implementation

The adaptive control equations, linear microdynamic controller equations, and reference model equations were converted from the continuous-time (Laplace) domain to the discrete-time domain via the bilinear approximation. Computer simulations written in ACSL (Advanced Continuous Simulation Language) were used to check the resulting difference equations. The computer simulation also showed that a digital controller sampling rate of 2 kHz was sufficient. The control scheme was implemented on a DSP-based hierarchical architecture [7].

Experimental results

Figures 6 and 7 show the response of the plant to 25 nm and 400 nm steps while under adaptive control. In this case the linear

microdynamic controller was tuned to give the reference model a slightly overdamped response with a rise time of approximately 0.1 seconds. Note that Figures 6 and 7 display very similar responses, regardless of the magnitude of the commanded step.

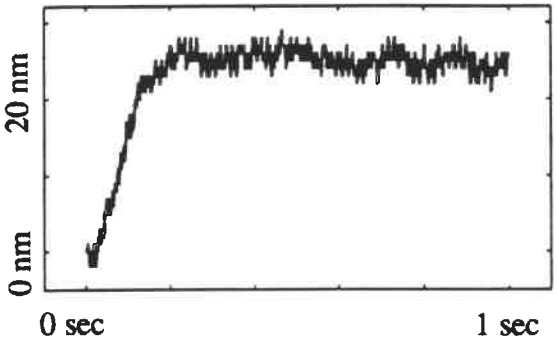


Figure 6: Response of plant ot 25 nm step command while under adaptive control.

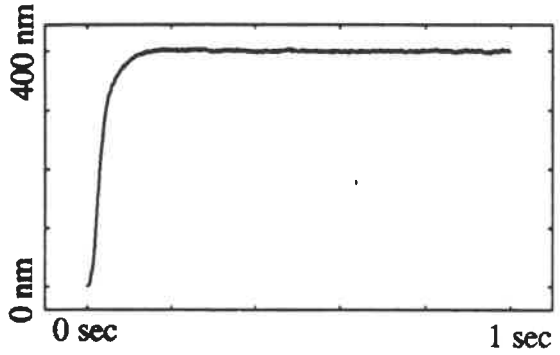


Figure 7: Response of plant to 400nm step command while under adaptive control.

Conclusion and Future Work

The model reference adaptive control approach has been applied to the microdynamic (pre friction breakaway) mode of the dual-mode dynamics observed experimentally. Experimental investigations have shown the need for a controller which can provide a magnitude independent (linear) closed-loop response. An MRAC scheme is applied and shown to be more robust with respect to the nonlinear dynamics than a PI scheme.

Future tasks include application of another MRAC scheme to the macrodynamic (post friction breakaway) mode of the plant. Computer simulation has shown the ability of such a scheme to compensate for the sliding friction torque which dominates large scale (micrometer) motion.

References

[1] Ro, P.I., P.I. Hubbel, "Nonlinear Microdynamic Behavior of a Ball Screw Driven Slide", submitted to *Precision Engineering*, 1991.

[2] Cuttino, J.F, T.A. Dow, "Nanometer Actuation Using a Precision Ball Screw", *Proceedings of the 1991 ASPE Annual Conference*, ASPE, 1991.

[3] Futami, S., A. Furutani and S. Yoshida, "Nanometer Positioning and its Microdynamics", *Nanotechnology*, Vol. 1, No. 1, July 1990, pp. 31 -37.

[4] Landau, Y.D., *Adaptive Control, The Model Reference Approach,* Marcel Dekker, New York, 1979.

[5] Gilbart, J.W. and Winston, G.C., "Adaptive Compensation for an Optical Tracking Telescope", *Automatica*, Vol. 10, pp. 125-131, 1974.

[6] Gilbart, J.W., R.V. Monopoli and C.F. Price, "Improved Convergence and Increased Flexibility in the Design of Model Reference Adaptive Control Systems", *Proceedings of the 1970 IEEE Symposium on Adaptive Processes (9th) - Decision and Control*, IEEE, New York, 1970.

[7] Fornaro, R.J., K.P. Garrard and L.W. Taylor, "Architectures and Algorithms for Computer Control of High Precision Machine Tools", *Proceedings from ASPE Annual Conference*, pp.31-34, September 1990.

ALUMINUM OPTICS, BUT OF WHICH ALLOY AND WHAT TEMPER?

Robert S. Polvani and Chris J. Evans
Precision Engineering Division
Manufacturing Engineering Laboratory
National Institute of Standards and Technology

Compromising figure and finish against fabrication costs, optics are frequently commodity products. The practice of single point diamond turning - rather than polishing - to finish optics, and using aluminum as the construction material is a well recognized way of containing costs. Aluminum alloys are a low cost family of engineering metals offering a combination of useful properties, good machinability, high thermal diffusivity and specific strength. However, they are particle hardened, and have an unstable microstructure. These two are serious limitations to a part's finish and figure, and so they act to its control production costs.

The particles cause poor finish and can cause figure problems through tool chipping and excessive abrasive wear. Microstructure instability means an optic will change figure with time. The question you might ask is need these be inherent shortcomings of aluminum? We hoped not, and have tried to identify an alloy and temper that would at least minimize the problem. Pure aluminum is a soft low strength material. For engineering uses it must be strengthened, and this is the source of the our problem.

While the alloys may appear to be strengthened in a variety of different ways, there is a common goal, and which is to reduce the mean free path of the mobile dislocation density. Chemistry is an effective widely used means. Alloying is used to warp or spring the lattice planes which hinder simple dislocation glide, to form impurity atmospheres about the dislocations which must then be dragged through the lattice, or to form obstructive particles which must be sheared or looped about. Formation of a dispersion - a regular network of coherent particles with nearest neighbors separated by about 0.1 micron - is the most effective form of chemical strengthening.

Materials processing is another means of strengthening. Forging reduces the mean free path in two ways. It creates more still dislocations to interfere with one another; and it refines the grain size. Both reduce mobility. The disadvantage of forging is dislocations are highly unstable structural defects. Readily interacting with one, they either form new substructure features or annihilate themselves. The commercial alloys reflect a cocktail of chemical and processing strategies, which use tempering to tailor the microstructure to the needs of a particular application.

Typically, optics use 6061-T6 or T651 aluminum, and a small usage made of cast alloys, 203 and A356, or other wrought alloys, 1100, 2024-T351, 5083, 5086. The TX51 tempers designate mechanically stress relieved stocks. Unless super purity aluminum is used, iron and silicon are always found in aluminum; and because iron is insoluble, phases of aluminum-iron or aluminum-iron-silicon will be found. The 1000 series aluminum are essentially pure, and for this reason they are soft low strength construction materials. The 2000 and 6000 series aluminum are age hardened materials. 6000 series uses the Mg_2Si intermetallic for the primary strengthening particle, along with the Fe_3SiAl_2 and $Fe_2Si_2Al_9$ intermetallics. But only the magnesium intermetallic responds to tempering treatments. The others form as the cast-ingot cools. The 2000 aluminum have a complex chemistry which includes copper, magnesium, manganese and the ever present tramp iron and silicon concentrations. Tempering transforms the iron phases to Al_7Cu_2Fe, the manganese forms a dispersion of $Cu_2Mn_3Al_{20}$ particles, but large residual particles of undissolved Al_2CuMg and $CuAl_2$ remain. The 5000 series are strain hardened aluminum. Magnesium and low temperature annealing is used to form precipitates of Mg_2Al_3. Chromium, a frequent additive, likely appears as a dispersion of Cr_2MgAl_{18}. If manganese is present the iron rich phases become complex.

The large inclusions in commercial aluminum must be properly appreciated. They are not simply dirt, but inherent and traceable to the initial cooling of the ingot. They are based not only on the iron and silicon impurities, but also heterogeneous nucleation and then excessive growth of the working dispersoids. Viewed chemically, the 5000 series alloys should be relatively clean stock. Our metallographic observations of 2024-T351, 5086-H112 and 6061-T351 samples, show little difference among them. All had roughly equivalent densities of large inclusions that were traceable to the ingot. It is these large particles that lead to excessive tool wear and pull-out scratches. They trouble single point diamond turning no less than polishing. For this reason aluminum optics are frequently single point diamond turned, electroless nickel plated, and then either single point diamond turned again or polished, or if necessary both.

Since wrought aluminum are either rolled or extruded, they are strongly textured and the strengthening particles highly stringered. We try to use this feature to good advantage. Texture means the grains are aligned along the working direction, have a high aspect ratio and a single strongly preferred crystal direction. A comparison can be made to the traditional corn broom. When possible, we machine the optic so that we cut across the whisker-like grains and particles. Heat treating is also used to reduce the particle size. The standard T6 tempering schedule is used except with a extended solution time. About 2 hours at temperature per inch thickness is used. The extension has two beneficial effects, it significantly reduces the cross sectional size of the particles and leads to pinching off of the stringers. A more spherical shaped particle results. The size gives readily shearable particles, and the shape rollable particles. While some

few tumble ahead into the surface, most roll out. In a 6061-T6 aluminum pull out scratches were largely eliminated. No negative evidence - grain growth, or eutectic formation at the grain boundaries - was found. No attempt was made to quantify the tool wear.

If finish were the only requirement, aluminum would be used in the hardest tempers. High hardness minimizes finish problems and aids machinability. Consider, 2024-T351 is harder than 6061-T651, which is harder than 5086-H112; the HVN are 150, 110 and 80 respectively. Their respective ASM machinability ratings are B, C and D.

Figure stability requires using soft tempered, stress free aluminum. The dimensional instability that follows unbalanced stresses or stress relaxation is obvious. Using X-ray diffraction, we found our 6061-T651 stock to be stress free in comparison to the 2024-T351 and 5086 samples, but the 6061 was also more sensitive to machining damage. For the 2024-T351 and 6061-T6, the recommended anneal - 190° C for 1-2 hours - was ineffective, there was no reduction of machining stresses. Instead, temperatures of 300° C or higher were necessary to reduce either conventional or single point diamond machining damages. The most effective way to relieve fabrication stresses is to redo the tempering on the semi finished blank, using a boiling water quench. A slow quench is necessary to minimize part warpage. To improve on the boiling water we will test polyalkylene glycol or polymer quenching. The claims made for the polymer are less part distortion than boiling water even though the hardness is equivalent to an iced brine quench. A last possibility for reducing quenching distortion is "uphill" quenching. The part is quenched into a dry ice and alcohol mixture and then immediately heated to about 190° C with steam. The principle is to use the thermal penetration lag between the interior and surface to distribute the stress throughout the entire piece, which serves to reduce the surface stresses by raising the internal stress.

The need for soft tempered aluminum may be less clear. Unfortunately, hard tempered alloys are high energy systems that revert to their lowest level. The process involves converting a population of many small particles into one with a few large members. The energy driver is reduction of particle surface area, which is accomplished by diffusion. This process is inevitable and highly temperature dependent. It will occur over several days at room temperature - the naturally aged tempers, and over hours at above 150° C. Depending on the alloy, the conversion affects figure because it causes parts to expand or contract. The particles have not only a different chemistry but a specific volume. To avoid dimensional instability problems ASM recommends overaging the alloys, or using the T8 temper for 6061 or 2024. Currently we are attempting to evaluate the long term dimensional instability of the 6061-T8 and annealed 5086-H112 tempers.

Single Point Diamond Machining Induced Damage on Single Crystal Copper, Aluminum and Gold.

Jeffrey W. Carr and C. Narayan

IBM Research Division
Thomas J. Watson Research Center
Yorktown Heights,N.Y.,10598,USA

J. Kim

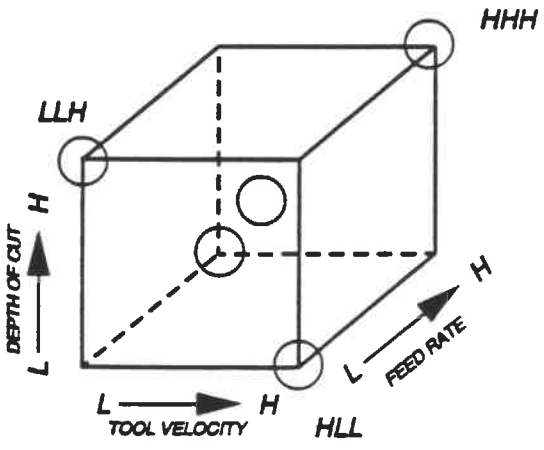
Goldstar Electronics
16 Woomyeon-Dong
Seocho-Gu, Seoul Korea

ABSTRACT

We have attempted to measure the damage induced in some metals by single point diamond flycutting. A fraction of the energy expended in machining causes dislocations in the metal below the cutting plane. Essentially, the dislocations rearrange the subsurface structure into a large number of smaller crystals with various orientations. Typically, the grains along the milled surface are smallest in size, as little as 10 nm. The grain size gradually increases with increasing distance away from the surface until a final grain boundary is reached, beyond which is the remainder of the single crystal.

The extent of damage is measured by observing a machined sample in cross-section with a transmission electron microscope (TEM). The limit of damage is taken to be the interface between the last grain boundary and the bulk of the remaining single crystal. This limit can be assumed to represent a minimum distance for damage. Undoubtedly, a number of dislocations at a density lower than that required for grain formation exist in the single crystal beyond the grain boundary. Unfortunately, we are unable to detect these features with the TEM.

The depth of damage for the {100}, {110} and {111} crystal orientations of copper and aluminum was measured for a number of different operating conditions. Specifically, we investigated the effect of tool velocity, feed rate and depth of cut. The determination of critical and non- critical variables allowed a reduction in the scope of the experiments on the relatively expensive gold crystals. Copper and aluminum samples consisted of small disks varying in diameter from 9.5 mm to 12.6 mm and approximately 1 mm in thickness. Initial samples were received with polished surfaces from the supplier (Monocrystals Company, Cleveland Ohio). Later experiments used disks sliced by wire EDM from single crystal rod.



Sample preparation for TEM is a tedious and time consuming endeavor. In order to maximize the data for the time invested we chose a multivariant approach to experimental design (1). Essentially, this approach permitted the investigation of three operating parameters; tool velocity, feed rate and depth of cut, with a reduced number of experiments while still covering the entire operating range of our present tool. Currently, we are able to adjust feed rate between 1 micron / rev. and 200 microns / rev.. To explore this range in the simplest possible manner, two values are chosen between the two extremes. For this experiment we investigated the damage at 10 and 120 microns / rev.. Similarly, tool velocity was set at a high of 30.7 m/s and a low of 15.3 m/s; depth of cut ranged from 0.3 microns to 6.0 microns.

The three variables can be thought of as three dimensions in an experimental space (figure 1). Each variable is an axis with a high end and a low end. Traditional experimental design would call for a data point at each corner of the cube for a total of eight experiments. Of course, to reduce the effect of a bad data point, each experiment would have to be performed three times, creating 24 samples for TEM analysis. A matrix must be executed for each crystal structure (there are three of these) and for each metal (again there are three). The total number of experiments has now increased to 216. The preparation and cross sectioning of these samples would take a competent technician the better part of a year.

Fortunately, the number of experiments can be drastically reduced by changing more than one variable at a time. As long as the number of experiments is at least as large as the number of variables, the required information can be reduced from the experimental data with simple linear algebra. The necessary experiments are highlighted by the four circles on the corners of the experimental space (figure 1). We have reduced the number of experiments from eight to four. Furthermore, multiple experiments at each experimental site can be avoided by measuring the noise of the system at some average condition. For this experiment, we chose a point in the center of the experimental space; representing a tool speed of 23 m/s, a 3 micron depth of cut and a feed rate of 60 microns / rev.. This experiment is performed three to five times and the spread of the data is taken as a measure of noise or uncertainty in the system. This uncertainty is then applied to all other experiments to determine if a measured difference is statistically significant. At this point, we have seven experiments for each crystal structure instead of 24. For three metals and three crystal structures, the required number of experiments would be 63 rather than 216.

The sample matrix could be reduced substantially from 63 by including crystal structure and metal type as two more variables. Investigations would then be designed around a five dimensional experimental space. Unfortunately, in our case, the uncertainty of sample supply made that approach risky. Diamond tool geometry, possibly another important variable, was not included in the first series of experiments. All operation were conducted with a 100 mm radius tool with a 0° rake face and a clearance angle of 8° (ETI, Indianapolis, IN)

A substantial amount of damage is induced into the crystal by the EDM process. To reduce the effect of this procedure on the final damage layer, each surface was prepared under conditions to limit damage. Initially, the samples were chemically polished to remove most, if not all, of the damage. Finally, a machining protocol with a slow feed rate and a very shallow depth of cut was developed to keep damage to a minimum. After the surface had been prepared, a single cut at the predetermined experimental condition was taken. It was possible to get a qualitative measure of residual subsurface damage by obtaining an electron diffraction pattern from the surface. A good surface will have a sharp pattern with a limited number of spots and a damaged surface will have a greater number of spots that appear poorly focused. While only semiquantitative at best, this method gives a simple, nondestructive measure of damage.

Of the three metals investigated, we have studied the damage in copper most extensively. A typical cross section with the corresponding diffraction patterns is seen in figure 2. Before milling we find a clean, undisturbed subsurface region. After milling, the number of diffraction spots increases substantially while becoming more diffuse. In the cross sectional view the number of new grain boundaries is apparent. The 100 nm bar gives an indication of size. At four microns below the surface some residual damage can still be observed for the particular conditions of this experiment; however, the overall damage is substantially reduced from the surface region. The cross section itself appears to be free from grain boundaries. The increase in scatter in the diffraction pattern is likely a result of dislocations beyond the last observed grain boundary.

In copper, damage ranges from a maximum of 4.4 microns to a minimum of 2.3 microns. The depth of damage was found to be remarkably unaffected by the tool velocity and the depth of cut. Essentially, there was no effect within the error of the experiment (+/- 0.2 micron). The feed rate was observed to have a measurable effect with the greatest damage occurring at the highest feed rate. For {100} copper, the damage depth increased from 3.7 microns to 4.4 microns when the feed rate was increased from 10 to 120 microns / rev.

The greatest influence on depth of damage arises from the orientation of the crystal structure. The damage extends 4.4 microns below the surface when a {100} surface is machined. However, the

depth of damage is reduced to 3.4 microns for a {110} plane and to 2.3 microns for a {111} plane. The difference can be understood empirically with a simple understanding of crystal structure (2,3). All metals that are considered compatible with diamond machining have a face centered cubic orientation (FCC). Included in this group, besides copper and aluminum is gold, nickel, lead, iridium, cobalt, rhodium, palladium, platinum and silver (as well as a few more exotic things). FCC is the easiest structure to deform. Twelve slip systems are present and all are active at the same time. In addition, all of the slip directions are along the {111} plane. That is the reason that subsurface damage is least when the {111} plane is parallel to the cutting plane. Dislocations occur along the surface rather than down into the sample. The other orientations are tilted with respect to the slip plane. A {110} surface, which shows moderate damage, is tilted 35.3°to the {111} plane. The force induced by machining causes dislocations along the slip plane now directed down into the sample, increasing the depth of the damage. For a {100} orientation, the slip plane is 54.7°from the cutting plane. The increased angle results in a greater depth of damage.

Aluminum is also FCC but on first inspection appears to behave differently than copper with respect to crystal orientation. The difference may be a result of the change in modulus between the two crystal orientations for the two metals. Unfortunately, due to the malleable nature of gold, sample preparation is difficult and damage measurements are inconclusive at this time. However, we observe that both materials are effected by tool velocity, possibly for different reasons. Superior surface finish was achieved on aluminum when the tool velocity was increased. The opposite effect was noted for gold; a better finish was obtained at lower speeds. Surface finish, measured with a Digital AFM/STM was used in conjunction with cross sectional TEM as a measure of damage. As in the case of copper, feed rate was also an important variable for both gold and aluminum.

Aluminum samples showed induced crystal growth to 3.2 microns below the surface in the {111} orientation; an increase over copper of 0.9 microns. In the {111} direction the modulus for aluminum is less than 40 % the value of copper. The decrease in modulus leads to an increase in deformation for a given force and a resultant increase in damage. In the {100} the modulus for the two materials is nearly the same. This fact is reflected in the depth of damage; 4.2 and 4.4 microns for aluminum and copper respectively

The change in tool velocity required for the successful machining of aluminum and gold may also be a explained by a difference in material properties. The melting temperature of aluminum is substantially lower than copper. As in the case of polymers, an increase in velocity means an increase in the temperature of the cutting zone. The melting temperature of aluminum (660°C for aluminum vs 1064°C for copper) may be low enough for the metal to be effected by the heat generated under the conditions of the experiment. The melting points for gold and copper are nearly identical so we must look elsewhere for an explanation of the low speeds required by gold. The gold surface most extensively investigated had {100} orientation. The modulus in the {100} direction is less than 65% that of copper. The increased ease of deformation can account for the decreased surface quality that is observed on gold.

The effect of damage is likely to be somewhat dependant on individual machines and diamonds. However, the trends in damage as a function of operating parameters and crystal orientation are expected to be independent of specific tooling. In bulk materials with all possible orientations, the damage is expected to be somewhat less than the maximum damage observed with the {100} structure. We are presently investigating the effect of machining on copper in polycrystaline form.

References

1) Practical Experimental Designs for Engineers and Scientist, William J. Diamond, Van Nostrand Reinhold Company, New York, 1981.

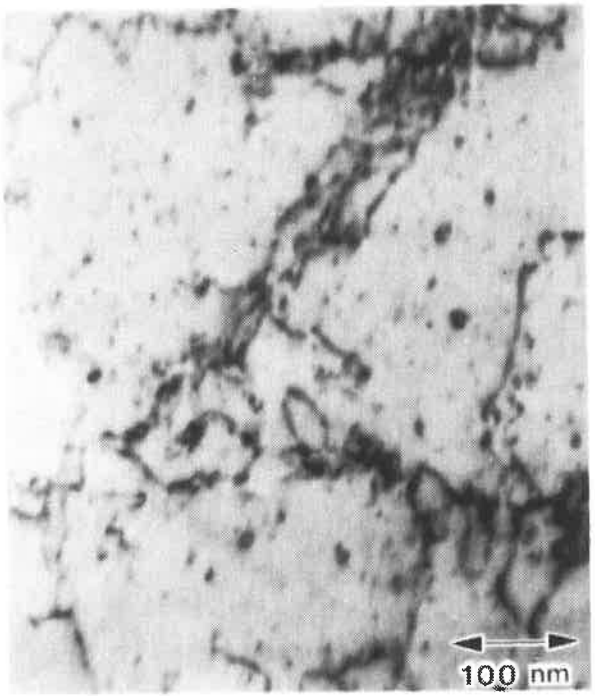
2) Deformation and Fracture Mechanisms of Engineering Materials, Richard W. Hertzberg, John Wiley and Sons, New York, 1979.

3) Elements of X-ray Diffraction, 2nd ed., B. D. Cullity, Addison- Wesley, Reading Mass., 1978.

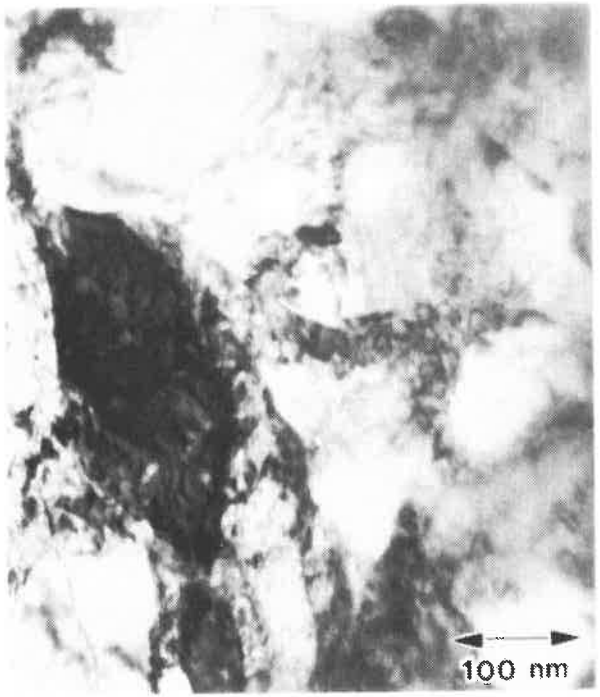
[MICRO – MILLING]

Cross-Section TEM of Single Crystal Copper

Before Milling



After / Near the Milled Surface



After / 4 μm Below the Surface

Molecular Dynamics Simulation of the Chip Forming Process in Single Crystal Copper and Comparison with Experimental Data*

I.F. Stowers, J.F. Belak, Lawrence Livermore National Laboratory; D.A. Lucca, R. Komanduri, Oklahoma State University; R.L. Rhorer, Los Alamos National Laboratory; T. Moriwaki, Kobe University; K. Okuda, Kobe Technical College; N. Ikawa, S. Shimada, H. Tanaka, Osaka University; T.A. Dow, J.D. Drescher; North Carolina State University.

Background: The precision machining of metal surfaces and the precision grinding of glass and ceramics are similar in that both can be modeled by the orthogonal cutting process. Despite the ability to fabricate these components, little is understood of the basic atomistic mechanisms by which the material is deformed and removed, by which the tool-tip and workpiece interact, and by which induced surface and subsurface damage occurs. It is for these reasons and because fracture and plastic flow are inherently atomistic in nature rather than continuous as assumed by continuum mechanics, that we are modeling deformation, using realistic many-body atomistic interaction potentials and state-of-the-art computational techniques.

Molecular Dynamics Modeling: Recently it has become possible to construct large scale (50,000 - 1,000,000 atom) molecular dynamic (MD) computer simulations of the indentation and orthogonal cutting of simple metals.[1,2,3,4,5] We have constructed such a simulation for the steady-state orthogonal cutting of single crystal copper. The model reveals considerable information on the evolution of chip morphology, stress fields and temperature distribution in the deformation zone, and the effect of tool sharpness (edge radius) on cutting forces and specific energy. The ability to model the orthogonal cutting process with a depth of cut of 20 nm has allowed our cutting force and specific energy predictions to be directly compared with recently reported measurements taken during ultra-precision machining experiments with single-point diamond tools. All of the experimental and simulation data reported in this paper were done with a 0° tool rake angle.

The MD simulations were done assuming a 2-dimensional embedded-atom (EA) type metal and scaled to the properties of copper. The simulated material was scaled to the properties of copper using the lattice spacing, cohesive energy, and atomic weight. The simulations were performed at 1/4 of the absolute melting temperature (66°C for copper). The EA type potential is an extension of the two-body model for metals and realistically accounts for the effect of the free-electron gas that surrounds each atom. It is a more realistic potential function and correctly models the change in properties of metals near a free surface. The EA model is however, limited to close packing materials like copper because it does not predict any angular dependency in the bonding and therefore will not correctly handle cubic and hexagonal metals.

Orthogonal cutting is the most general simplification of the metal cutting process and assumes that the material is sufficiently wide such that edge effects have little or no influence on results (plane-strain condition). In general, the center line of the chip is assumed to be a plane of symmetry. Our computer modeling was done with only one layer of material to both simulate this symmetry plane and to obtain nanometer length scales in reasonable computer times. Atomistic motion outside of this one layer is not allowed as in the center line of the orthogonal cutting process. However, this is not the most general

---

[1] W.G. Hoover, C.G. Hoover, I.F. Stowers, W.J. Siekhaus, **Interface Tribology by Nonequilibrium Molecular Dynamics,** *Mat. Res. Soc. Symp.*, Vol 140, Boston, MS, 1989. pp. 119-124.

[2] W.G. Hoover, C.G. Hoover, I.F. Stowers, A.J. De Groot, B. Moran, **Simulation of Mechanical Deformation via Nonequilibrium Molecular Dynamics,** *CECAM Workshop on Microscopic Simulation of Macroscopic Flows*, Brussels, Belgium, Aug 1989.

[3] W.G. Hoover, A.J. De Groot, C.G. Hoover, I.F. Stowers, T. Kawai, B.L. Holian, T. Boku, S. Ihara, J. Belak, **Large-Scale Elastic-Plastic Indentation Simulations *via* Nonequilibrium Molecular Dynamics,** *Physical Review A*, 10 (42), 1990, pp. 5844-5853.

[4] J.F. Belak, I.F. Stowers, **Molecular Dynamics Studies of Surface Indentation in Two Dimensions,** *Atomic Scale Calculations of Structure in Materials*, Vol 193, San Francisco, CA, Apr 16-21, 1990. pp. 259-264. Materials Research Society.

[5] J.F. Belak, I.F. Stowers, **A Molecular Dynamics Model of the Orthogonal Cutting Process,** *ASPE Annual Conference*, Rochester, NY, Sep 23-28, 1990. pp. 76-79. ASPE, PO Box 7918, Raleigh, NC 27695.

assumption and future simulations will be done with several layers of material to allow dislocation motions outside of the symmetry plane.

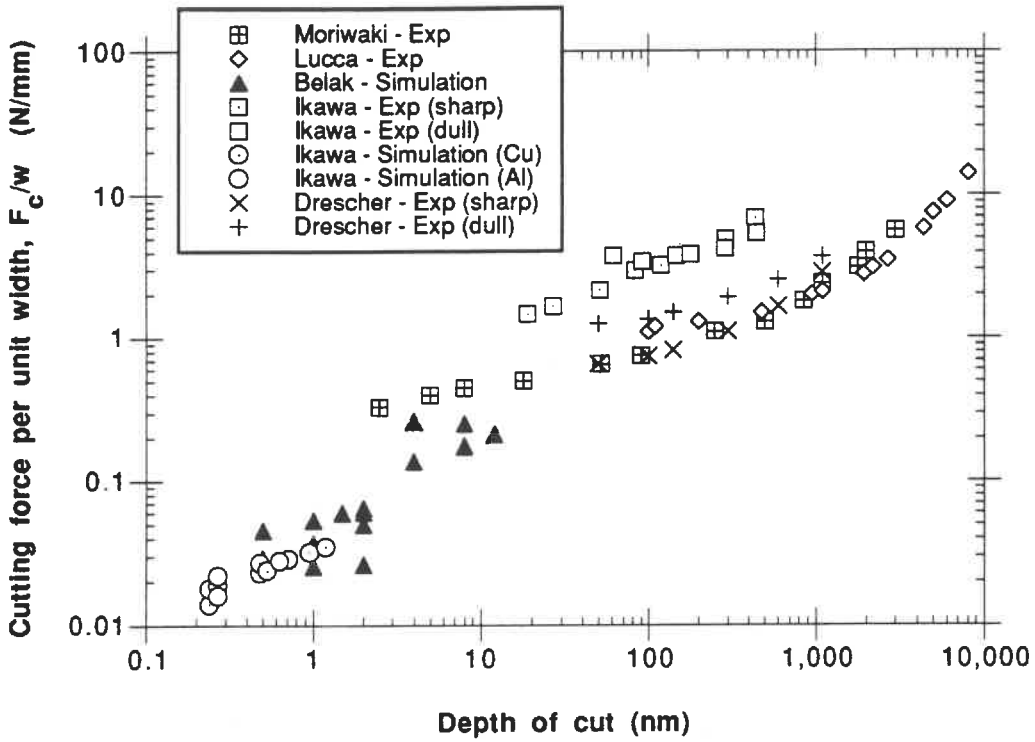


Figure 1 Cutting force per unit width in the horizontal direction (F_C/w) versus the depth of cut (d). All of the experimental data generally follow the same trend line and are clustered surprisingly close together considering the wide differences in tool edge radius and cutting speed. The molecular static simulation (Ikawa) predicts forces that appear to agree with the MD simulation (Belak) and form an extension to the experimental data at depths of cut of only a few Angstroms. In general, the cutting force is proportional to the square root of the depth of cut.

Experimental and Molecular Statics Data: The orthogonal cutting experimental data from Moriwaki[6] was obtained with OFHC polycrystalline copper with a 10-100 μm grain size, a 20-50 nm estimated tool edge radius, and was cut at 23 m/s. The orthogonal cutting experimental data from Lucca[7] was obtained with a similar OFHC copper, a 100-200 nm estimated tool edge radius, and cut at 0.1 m/s. The orthogonal cutting experimental data from Ikawa[8] was obtained with a similar polycrystalline copper, a 20 nm (sharp) and 2000 nm (dull) estimated tool edge radius, and cut at 7-17 m/s. The molecular statics (MS) computer simulation data from Ikawa[9] assumed a Morse Potential quasi-static type material being cut under near static

---

[6] T. Moriwaki, K. Okuda, **Machinability of Copper in Ultra-Precision Micro Diamond Cutting,** *Annals of CIRP*, Vol 38/1, Tronheim, Norway, pp. 115-118.

[7] D.A. Lucca, R.L. Rhorer, R. Komanduri, **Energy Dissipation in Ultra-Precision Machining,** *ASPE 1990 Annual Conference*, Rochester, NY, Sep 23-28, 1990. pp. 51-54. American Society for Precision Engineering, PO Box 7918, Raleigh, NC 27695.

[8] N. Ikawa, S. Shimada, H. Tanaka, **Minimum Thickness of Cut in Micro-machining,** *2nd Nanotechnology Seminar*, University of Warwick, Coventry, UK, Aug 23-24, 1990.

[9] N. Ikawa, S. Shimada, H. Tanaka, G. Ohmori, **An Atomic Analysis of Nanometric Chip Removal as Affected by Tool-Work Interaction in Diamond Turning,** *Annals of CIRP*, Vol 40/1, Stanford, USA, in publication.

velocity and temperature conditions. The orthogonal cutting data from Drescher[10,11] was obtained with an electroplated polycrystalline copper, a 50 nm (sharp) and 1000 nm (dull) estimated tool edge radius, and cut at 2-3 m/s.

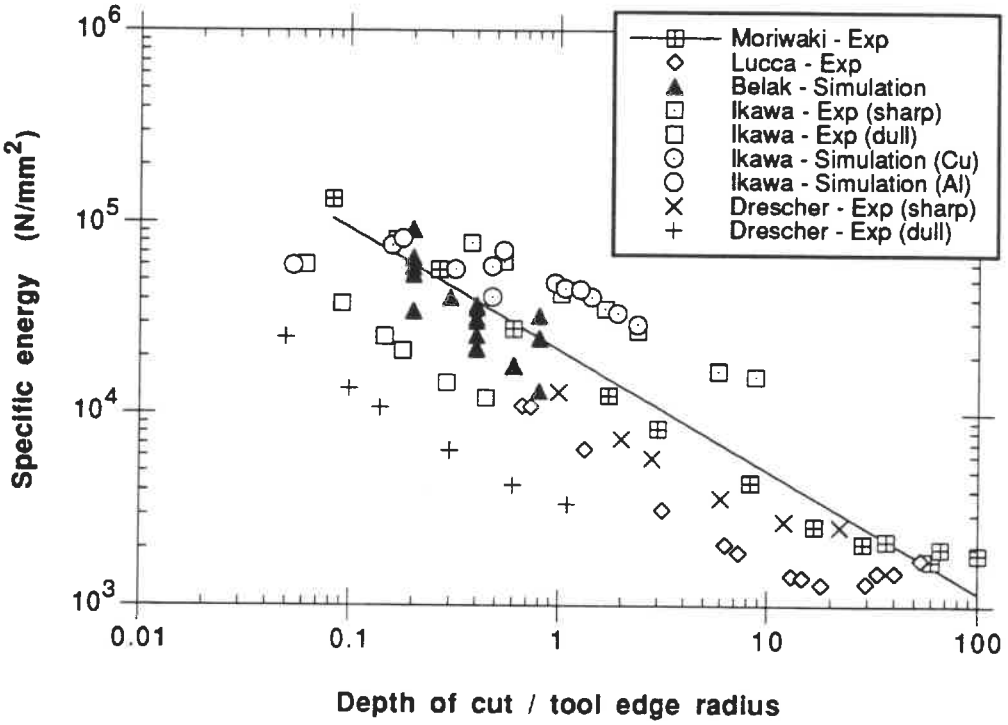


Figure 2 Specific energy (E_S) required to machine copper plotted against the depth of cut (d) normalized by the tool edge radius (R_e). Specific energy is the cutting force (F_C) multiplied by the length of the cut (l) divided by the volume removed (l * w * d) where w is the width of the cut. This reduces to $E_S = F_C / (w * d)$. There is good agreement between the experimental data and the MD and MS simulation data. The horizontal scatter in the experimental data may be due to a lack of knowledge of the exact value of the tool edge radius. A curve fitted through Moriwaki's data and parallel to the other data fits a curve described by: $E_S = 21645\ (d/R_e)^{-0.634}$.

Findings and Conclusions:

This data shown in Figure 2 reflects the classic "size effect" reported by Shaw[12,13] and Armarego[14]. Shaw reported, however, that the specific energy is proportional to $d^{-0.2}$ in the conventional machining chip thickness range. The specific energy consumed at the shallowest depths of cut asymptotically approaches a value slightly larger than the heat of formation of copper (4.8×10^4 N/mm^2). That is, there is sufficient

---

[10] J.D. Drescher, T.A. Dow, **Machining Forces in Diamond Turning of Plated Copper and Unplated Substrates**, *Proceedings of the ASPE Spring Topical Meeting: Metal Plating for Precision Finishing Operations*,, April 1991, American Society for Precision Engineering, PO Box 7918, Raleigh, NC 27695,USA.

[11] J.D. Drescher, T.A. Dow, **Tool Force Model Development for Diamond Turning**, *Precision Engineering*, 1 (12), 1990. pp. 29-35.

[12] W.R. Backer, E.R. Marshall, M.C. Shaw, **The Size Effect in Metal Cutting**, *Transactions of ASME*, (74), 1952, pp. 61-72.

[13] M.C. Shaw, **Metal Cutting Principles**, *Clarendon Press*, 1986, p. 42.

[14] E.J.A. Armarego, R.H. Brown, **On the Size Effect in Metal Cutting**, *Int. Jnl. Prod. Res*, Oct 25, 1961, pp. 75-99.

energy expended in forming the chip to transform the material from a solid to a vapor. It is still not clear whether melting and/or evaporation followed by rapid quenching actually occur, however, there is no question that sufficient energy to accomplish the transformation is available.

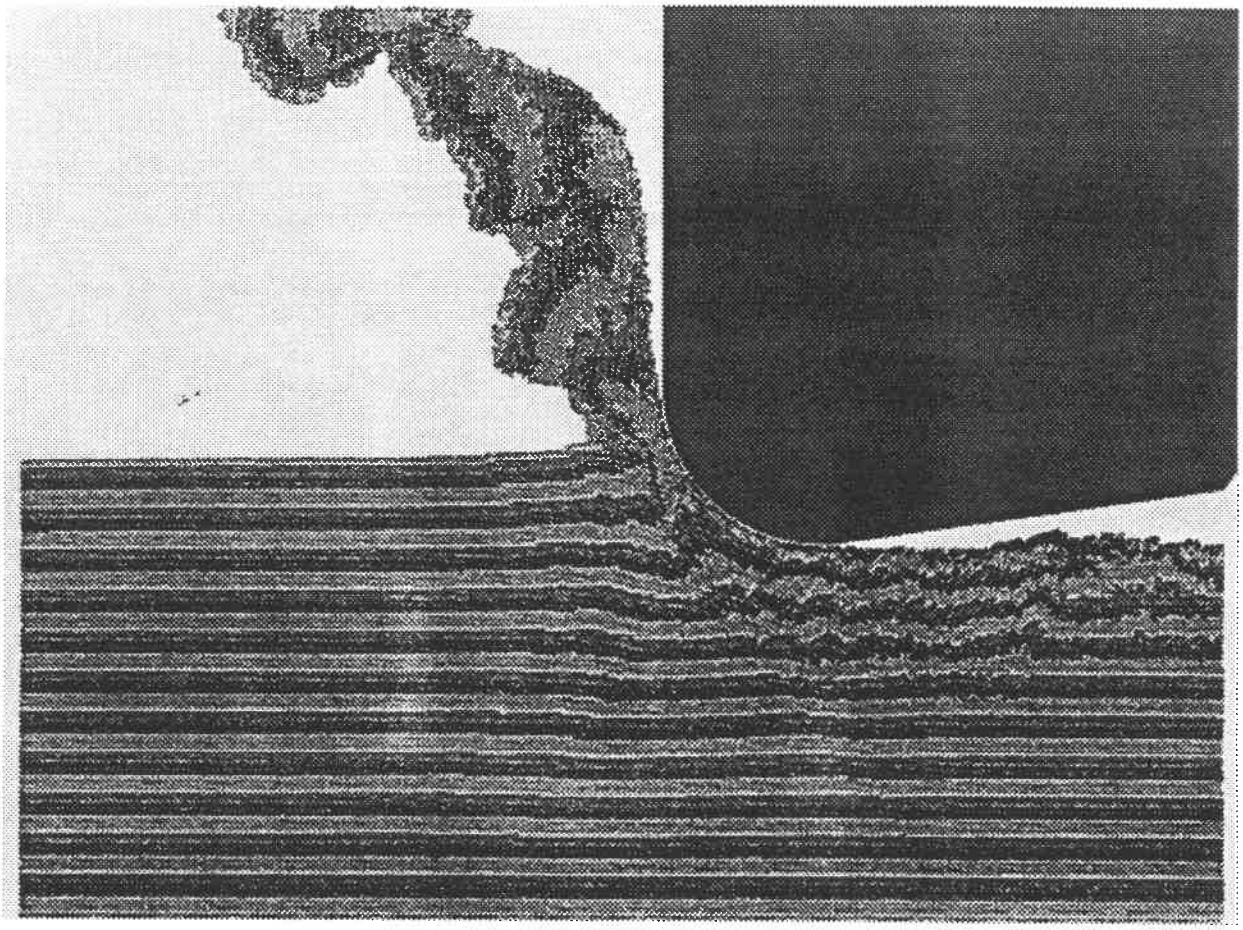


Figure 3 Molecular dynamics computer simulation of the chip formation process in copper with a round nose tool at room temperature with a depth of cut of 8 nm (32 atomic layers), tool edge radius of 20 nm, and a cutting speed of 86 m/s. This picture was generated after cutting for ≈ 1.5 ns or for a distance of 0.1 μm. The alternating regions of chip thickness variation occur at time intervals of ≈ 0.6 ns. The horizontal shaded bands in the figure represent intervals of 16 atomic layers for visualization purposes only.

In general, results from the molecular dynamics simulations agree remarkably well with the four sets of experimental data and the molecular static simulation data. Furthermore, the normalized cutting force is relatively independent of the cutting speed and the cutting force tends to decrease with a decrease in the tool edge radius (increase in sharpness) in qualitative agreement with the experimental data.

At the larger depths of cut the chip morphology is strikingly similar to that observed in actual metal cutting. The chip develops a sawtooth appearance as a result of alternating periods of chip thickness increase followed by rapid thinning. The chip forming process appears to be as follows: a shear band forms, the tool overtakes the shear band and it dissipates, and the shear band again reforms within the material. The chip formation process is therefore a quasi-steady state process rather than a steady state process.

**This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48.*

MACHINING FORCES IN DIAMOND TURNING OF COPPER

J. D. Drescher
T. A. Dow
North Carolina State University,
Precision Engineering Center, Raleigh, NC 27695-7918

Introduction

Good surface finish and dimensional accuracy are the two specifications which often require DT as a manufacturing process. The forces in DT must be understood. An understanding of how the tool forces change with cutting conditions, material, and tool wear allows the machine tool controller to maintain figure and surface accuracy using force as a feedback parameter.

Experiment

Measurement System Figure 1 is a schematic of the force measurement system. The main component is a piezoelectric, quartz transducer which is sensitive in three perpendicular directions. A preload is required so that transverse forces can be transmitted by friction. This preload is applied by a 6 mm bolt through the center of the transducer.

In the cutting force measurements described in this report, forces were measured in the two directions perpendicular to the preload axis. In this way the temperature effects on force were avoided. The directions of interest are labeled Y and Z in Figure 1.

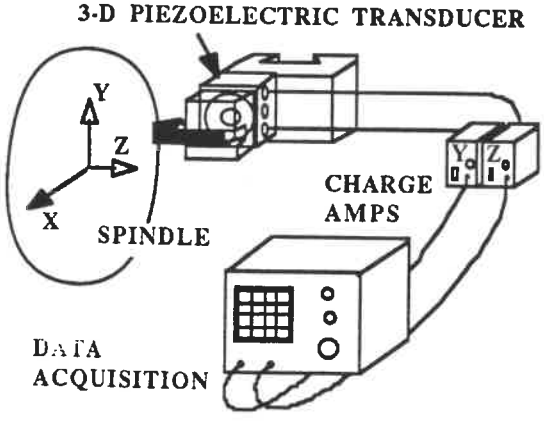


Figure 1: Force measurement system

Cutting Geometry The schematic in Figure 2 shows the orthogonal cutting geometry which is best for studying the turning process. The chip produced has uniform thickness and width. In this orthogonal geometry, the cutting parameters are limited to cutting depth and edge length. The feed is in the Z-direction so that feed/revolution is equal to cutting depth, d. The edge length, L, is determined by the length of tool employed in the cut. Tool rake angle is not a parameter of interest. 0° rake and 6° clearance angle tools were used in the work reported here. A steady flood of light machine oil was directed at the tool edge to entrain the continuous chip and carry it away.

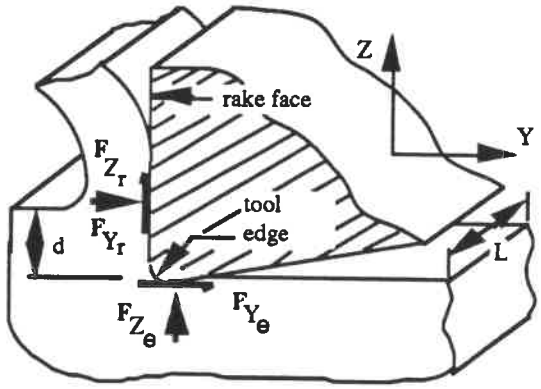


Figure 2: Orthogonal tool force model

Analysis

Tool Force Model There are four force components; two normal and tangential pairs. These can be added as vectors to describe the measured forces; the total forces in the cutting and thrust directions. The normal force on the rake face has been shown to be proportional to the area of the uncut chip, L d [1]. When depth of cut is small relative to edge sharpness, however, measured forces do not scale linearly with depth of cut. This is due to the forces on the tool edge. This pair of normal and frictional forces are not related to depth of cut.

They are not associated with material removal. The force model is:

$$F_Y = F_{Yr} + F_{Ye} = C_1 dL + \mu C_2 L \qquad (1)$$

$$F_Z = F_{Zr} + F_{Ze} = \mu C_1 dL + C_2 L \qquad (2)$$

There are three variables in this model, C_1, C_2, and μ. The friction coefficient, μ, is assumed equal on the rake and clearance face. With experimental force data, where d, L, F_Y, and F_Z are known, values of C_1, C_2, and μ can be found which optimize the correlation between measured forces and the model. Figure 3 shows measured tool force data (points) in the Y-and Z- directions as a function of the depth of cut. The best fit lines are also plotted using Equations (1) and (2). The values for the friction coefficient (μ= 0.21), and the parameters (C_1 and C_2) are included in the graph as well as the correlation coefficient of the data fit. A coefficient of 1.0 would correspond to a perfect fit. The coefficient C_1 represents a flow stress and has units of pressure (GPa) while C_2 has units of force/unit width (N/mm).

In the force model, the friction coefficient is the only one of the three parameters whose physical significance is readily apparent. However, the units on C_1 and C_2 give insight into their physical significance. C_2 (force/unit width) when multiplied by edge length, L, gives the normal force on the tool edge in the thrust direction. The area of the tool in this direction is A_2 = wL so that the stress on this area is C_2/w. C_1 is the stress acting on the rake face area, A_1 = Ld.

Wear Effects on Tool Force The forces on the tool change as a result of wear on the edge. The model described by Equations (1) and (2) can be used to indicate the source of these changes. As the model is fit to force data at different wear conditions, C_1 and μ do not change. By comparing the variation in C_2 with changes in the shape of the tool as measured using a high-resolution SEM, the tool condition can be incorporated into the model.

Figure 4 is a schematic of a tool on which a flat area has been worn on the flank face. Assuming that the stress is the same on the rake face and clearance face, the wear width, w, can be predicted based on measured forces. With this assumption, the relationship between the parameters C_1 and C_2 and the width of the worn region can be calculated. The normal force on the area Lw can be written as C_1Lw. This is also equal to C_2L.as seen in Equation (2) Therefore:

$$w = C_2 / C_1 \qquad (3)$$

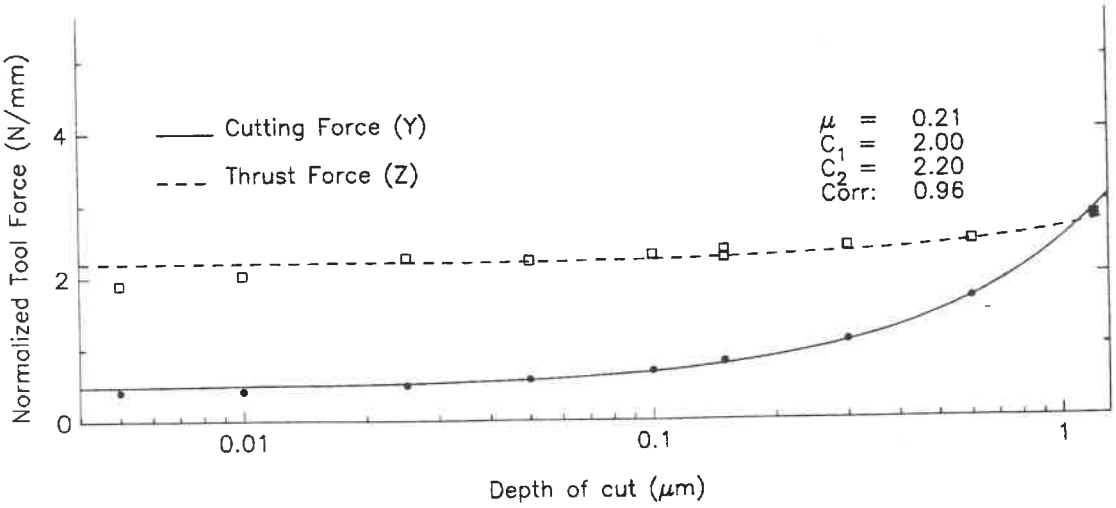


Figure 3: Curve fit for 3-parameter model

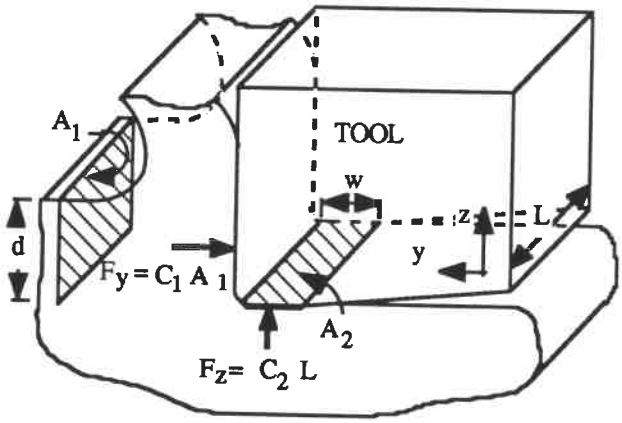


Figure 4: Tool edge diagram with edge wear

Equation 3 states that the flank wear can be calculated based on parameters derived from experimental force data. Equivalently, the flank wear is included in the model for tool forces. Equations (1) and (2) become:

$$F_Y = C_1 [L d + \mu L w]$$

$$F_Z = C_1 [\mu L d + L w]$$

The parameter C_2 in the model has been replaced by the wear width which has physical significance. Force data can be used to find the wear width which best fits the data.

Measured forces were used to predict the wear in two separate series of wear tests. In the first series, a straight-edge tool was used to cut plated copper for a total of 28 km. Periodically, force measurements were made over a range of cutting depths and the best fit parameters were found using the model of Equations (1) and (2). The calculated values of w are plotted as +'s in Figure 5. The width of the wear land as predicted by this procedure suggests that tool wear was on the order of 20nm. For this series of experiments, no measurements of the tool were attempted.

A second series of tests was performed using Al 6061-T6 as the workpiece. Much higher tool wear was observed. Force measurements were made using the same procedure as the first series of tests. That is, while aluminum was used to accelerate wear, forces were measured cutting the plated copper. The calculated values for w are plotted in Figure 5 as the open circles. Note that after 1 km of cutting the aluminum, predicted wear is greater than that predicted after 28 km of cutting the copper.

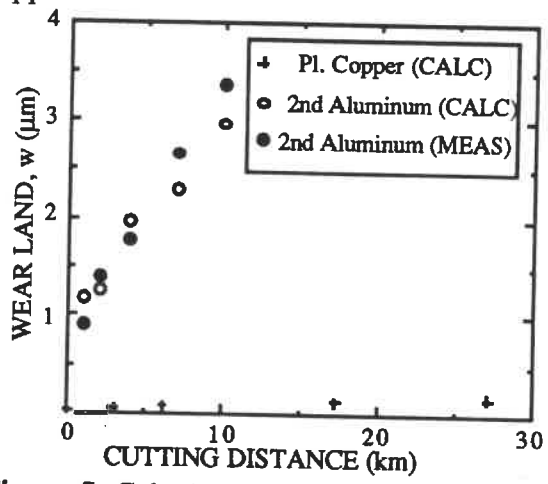


Figure 5: Calculated and measured tool wear vs. cutting distance

Figure 6: Micro-photograph of tool edge after 2km of cutting Al 6061-T6

At each stage of wear, the tool was examined in a field emission scanning electron microscope (SEM). The tool was imaged uncoated at 20,000X and 5 keV. One SEM photo used for the edge measurements is shown in Figure 6. Figure 6 is the photo taken of the tool after 2 km of cutting aluminum. Lapping marks are visible on both the clearance face and rake face of the tool. These have been worn away at the edge as the flat area developed. The wear dimension of interest is perpendicular to the rake face. Lines can be seen along the cutting direction on this wear area. The measured values of the wear width are plotted in Figure 5 as the filled circles. These values are in

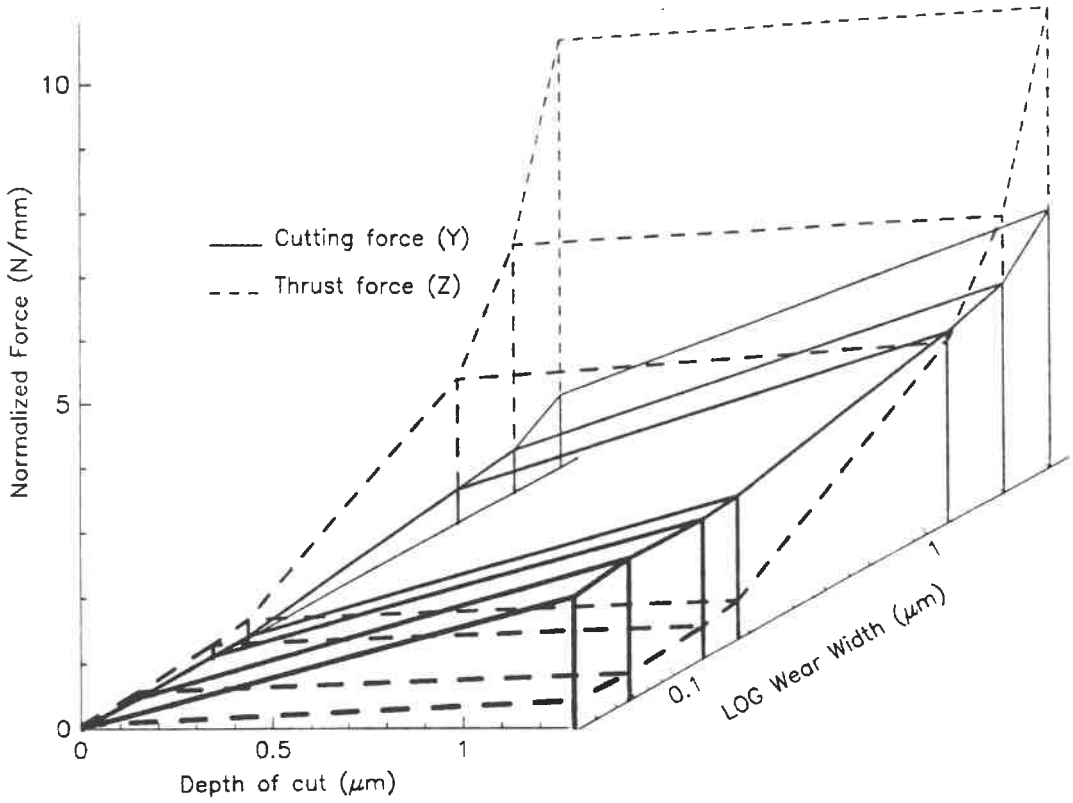


Figure 7: Tool force vs. edge wear and depth of cut using the model of Equations (1) and (2)

agreement with the wear predicted from the tool forces. The agreement of the measured and calculated tool wear indicates that using the force model to calculate wear width is a reliable procedure. The +'s in Figure 5, from the first test series, are thus assumed to be accurate.

Each calculated point in Figure 5 was obtained from force data similar to Figure 3. These measured forces can be combined into a single plot to show force vs. depth of cut and wear. The data covers nearly two decades of wear width from 41 nm to 3 μm.

Figure 7 is a plot of the normalized force components (Force/unit edge length) as functions of depth of cut and wear width. The cutting force is most influenced by depth of cut and to a lesser degree by tool wear while the thrust force is dominated by the tool edge condition. The ratio of measured thrust force to measured cutting force changes continuously along a line of constant cutting depth. Therefore, using the model represented by Figure 7, a single force measurement has enough information to assess the tool condition when depth of cut is known.

Discussion The terms flow stress and dynamic shear stress have been used interchangeably in metal cutting analyses. In metal cutting, flow stress is defined as the cutting forces resolved onto the shear plane divided by the shear plane area. Traditionally, the forces at the tool edge are considered negligible so that in terms of measured cutting force, F_c, and thrust force, F_t, the flow stress can be written:

$$\tau_s = \frac{F_s}{A_s} = \frac{F_c \cos\phi - F_t \sin\phi}{A_0 / \sin\phi} \qquad (4)$$

where,

ϕ = shear plane angle
F_s = shearing force on shear plane
A_s = shear plane area
A_0 = cross-sectional area of uncut chip

ANALYSIS AND MODELING OF DIAMOND TURNING OF BRITTLE MATERIALS

M. Tidwell and R. O. Scattergood
Precision Engineering Center
North Carolina State University
Raleigh, NC 27695-7918

INTRODUCTION

Research work at the Precision Engineering Center has been directed at gaining a more fundamental understanding of single-point diamond turning and diamond grinding of brittle materials. The material removal mechanisms in the so-called ductile regime are of particular interest in this context. Single-point diamond turning offers the best approach to fundamental studies because the machining geometry is easily characterized compared to grinding. The work reported here was carried out using single crystal germanium as a model material. The interrupted cutting test technique (ICT) developed in earlier work was modified and extended.

EXPERIMENTAL PROCEDURE

The ICT has been described earlier [1]. The tool is rapidly withdrawn from the workpiece during cutting and the damage transition on the uncut shoulder is examined. Figure 1 shows a projection along the uncut shoulder in schematic form. Fracture damage initiates at a cutting depth d_c at position z_{eff} from the tool centerline. The depth of damage is y_c. By careful analysis of the geometry shown in Fig. 1, taking into account the overlap of damage from successive tool cuts, one can derive the equations

$$d_{eff} = \frac{f}{R}\left(z_{eff} + \frac{f}{2}\right) \tag{1a}$$

$$d_{eff}^2 - \frac{f^2}{R}d_{eff} = d_c^2 - \frac{f^2}{R}(d_c + 2y_c) \tag{1b}$$

$$f_{max}^2 = \frac{Rd_c^2}{d_c + 2y_c} \tag{1c}$$

These contain corrections not included in previously published versions of the same formulae [1,2]. Measurements of z_{eff} vs. feed rate f for given cutting conditions and tool nose radius R allow the damage parameters d_c and y_c to be obtained experimentally from ICT data. f_{max} represents the maximum feed rate possible before fracture damage replicates into the finished surface, and thus represents a measure of machinability in the ductile regime. Interestingly, material removal by fracture can occur in the ductile regime, providing the damage shown in Fig. 1 does not replicate into the finished surface.

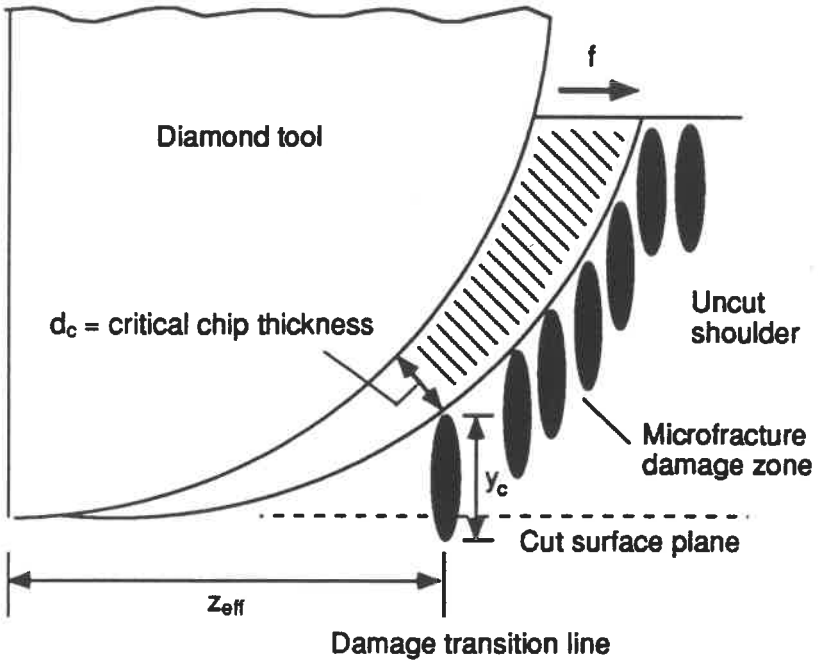


Figure 1. Schematic of an ICT viewed along the cutting direction (plastic zones shaded).

In order to facilitate the analysis of test data, and to obtain good statistics over a large area of damage zone, a computer-based image analysis technique was developed to analyze the fracture damage transitions on the uncut shoulder [1,3]. The density of microfracture is extracted from image analysis of optical (Nomarski) micrographs. Values of z_{eff} are obtained in conjunction with a data extrapolation technique for the zero damage limit . Machining tests were conducted using single-point diamond tools with R = 3.175 mm and rake angles of -10°, -30° and -45°. Single crystal germanium wafers with (100), (110) and (111) surface orientations were used for the tests. Multiple tests were made on a single wafer. All measurements were made for the cutting direction producing maximum damage on each surface orientation [4], for example, [110] on a (100) wafer. Further details on test procedure will be available elsewhere [2].

RESULTS AND DISCUSSION

Table 1 presents the results for the effect of tool rake angle and cutting coolant on d_c, y_c and f_{max} for (100) wafers. It is evident the the critical cutting depth d_c increases systematically as the rake angle becomes more negative. The fracture damage depth y_c also tends to increase with more negative rake, however, there is considerable variation in the values. Furthermore, the values of y_c are large compared to d_c. The net result, as obtained from Eq. (1c), is a significant reduction in f_{max} with increasing negative rake, an effect familiar to practitioners. It is important to note that while the increase in d_c with negative rake angle favors machinability, the benefit is negated to some extent by the variations in y_c. It is also evident from Table 1 that dry cutting (air) enhances ductile response relative to cutting using deionized water (wet).

Table 1. d_c, y_c and f_{max} as a function of rake angle and coolant for (100) surfaces.

Parameter	Rake Angle (deg)	Wet	Dry
y_c (nm)	-10	670	1484
	-30	2358	887
	-45	1889	1276
d_c (nm)	-10	40	57
	-30	119	181
	-45	312	270
f_{max} (μm/rev)	-10	1.5	1.8
	-30	2.8	7.0
	-45	8.2	8.7

Table 2. d_c, y_c and f_{max} as a function surface orientation and coolant at - 30° rake angle.

Parameter	Orientation	Wet	Dry
y_c (nm)	(100)	2358	887
	(111)	299	198
	(110)	550	178
d_c (nm)	(100)	119	181
	(111)	88	68
	(110)	140	270
f_{max} (μm/rev)	(100)	2.8	7.0
	(111)	6.2	5.6
	(110)	10.4	11.3

Table 2 presents the results for the effect of crystal surface orientation and cutting coolant on d_c, y_c and f_{max} using a - 30° rake angle. There is a significant and rather complicated effect of crystal orientation on machining response. In general, (110) surfaces have the best performance (largest f_{max}) followed by (111) and then (100). The dry cuts may overestimate performance to some extent since humidity was not controlled and often reached 60% (the wet cuts give a more standardized measure of relative machinability). The wide range of variation in the single crystal data is significant for applications using polycrystalline germanium. In the latter case, ductile-regime conditions are dictated by worst-case single-crystal limits. Textured polycrystal material would offer some advantages in this regard.

A preliminary interpretation of the results in Tables 1 and 2 has been obtained using a fracture-mechanics approach based on a simplified treatment of the cutting stresses [5]. A line force P applied to the surface of a half-space at an angle θ produces the elastic-plastic deformation field shown schematically in Fig. 2. Fractures will initiate in the region behind the tool. The critical

depth d_c follows from $K_{max} = K_c$ where K_c is the fracture toughness and K_{max} is the maximum stress intensity factor evaluated over all possible crack trajectories (mode I crack trajectories are orthogonal to the stress contours shown in Fig. 1). Finite-element analysis supports the predictions of this simplified model [6] for the smaller depths of cut appropriate to ductile regime condintions. At larger cutting depths, other fracture modes can appear. The adjustable parameters in the current analysis are the effective tool-workpiece friction coefficient μ, which in conjunction with rake angle determines θ, and the brittleness index $(K_c/H)^2$ where H is hardness. Values of μ and $(K_c/H)^2$ appropriate to germanium give good agreement for the d_c vs. rake angle results in Table 1. However, the y_c results are significantly underpredicted. This suggests that subcritical crack growth in the unloading (residual stress) field affects the fracture depth. Coolant likely plays a role in that environmentally assisted crack growth can occur during and after the unloading cycle. Crystal orientation effects arise due to anisotropy in the fracture surface energy (K_c is orientation dependent). Details of the latter response are complex and still under investigation.

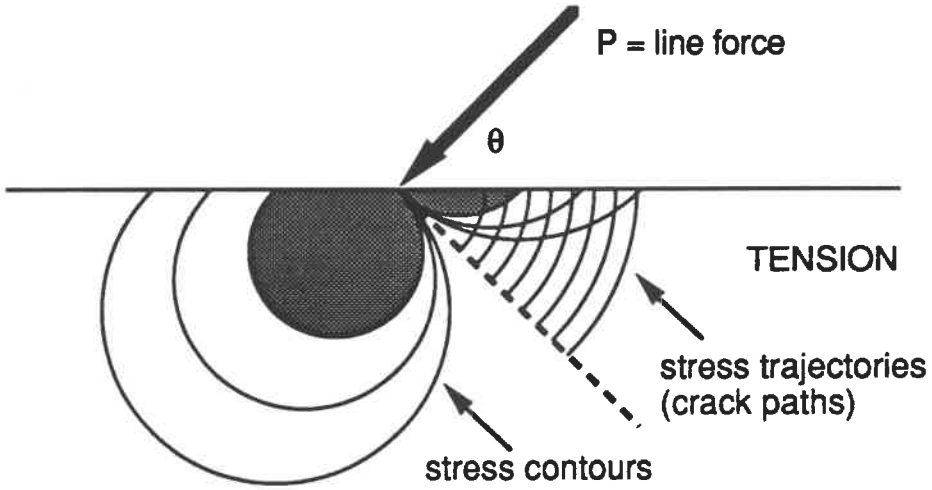


Figure 2. Schematic of the elastic-plastic stress field due to line force P.

REFERENCES

1. W. S. Blackley and R. O. Scattergood, "Ductile Regime Machining Model for Diamond Turning of Brittle Materials", Prec. Engr., vol. 13, no. 2, 95 (1991).
2. M. Tidwell, MS Thesis research in progress, North Carolina State University (1991).
3. M. Tidwell and R. O. Scattergood, "Diamond Turning of Brittle Materials", PEC Annual Report, vol. VIII, 153 (1990).
4. W. S. Blackley and R. O. Scattergood, "Crystal Orientation Dependence of Machining Damage - A Stress Model", J. Am. Ceram. Soc., 73, (10), 3113 (1990).
5. W. S. Blackley, M. Tidwell and R. O. Scattergood, in preparation (1991).
6. G. Hiatt and J. Strenkowski, "Fracture Mechanics Models for Predicting the Ductile Regime in Single Point Turning of Brittle Materials, PEC Annual Report, vol. VIII, 181 (1990).

Material Removal Mechanisms Involved in the Single-Point Diamond Turning of Brittle Materials

C-L Chao

Chung-Cheng Institute of Technology, Ta-Hsi, Tao-Yuan, Taiwan, R.O.C.

A.E. Gee

Cranfield Institute of Technology, Cranfield, Bedford, UK

1. Introduction

When subjected to increasing stress, glasses, ceramic materials and certain important hard crystals like silicon and gallium-arsenide, all fracture before appreciable plastic deformation takes place; this is (classically) said to be evidence of 'brittle' characteristics. Although such effects make the machining of these materials extremely difficult, the demand for precision parts made from them has risen (and continues to rise) primarily because of their superior physical, mechanical, optical or electronic properties. Advances in manufacturing processes are therefore important for the economic production of these parts.

For some years, it has been possible to diamond-turn, to mirror-like surface finishes, certain *traditionally-machinable* metals such as aluminium alloy, electroless nickel, and copper achieving sub-micrometre figure tolerances and surface roughness of under 10 nm Ra. However, when applied to brittle materials such as glass with cut depths of several micrometres or more, diamond-turning was seen to be a "crushing" process (Bryan and Carter 1985) , producing a surface characterised by conchoidal fractures and a damage-layer about 25 μm deep. Optical glass-workers and materials' scientists have found that glasses and other brittle materials do not necessarily manifest brittle fracture when machined with the finest cuts ($< 1\ \mu m$). Such phenomena have been found in grinding (Bifano et al 1988), single point diamond turning (Puttick et al 1989) and ruling (Chao and Gee 1989). It is apparent that, for cut depths below some critical value, these (macroscopically brittle) materials can be machined in a restricted *ductile-mode* whereby material is removed by plastic deformation processes rather than brittle fracture processes.

It is possible to foresee that ductile machining, analogous to that used on metals and capable of yielding an ultra-fine surface finish and low sub-surface damage on silicate glass, quartz and other brittle materials would present certain major advantages as a manufacturing process. This would open up the possibility for producing aspheric lenses in an economic and deterministic manner, reducing the cost and weight of multi-element (spheric) lenses and replacing the time consuming and indeterministic free-abrasive grinding and polishing operations traditionally used to produce surfaces on glass (spheric and plane).

In order to achieve ductile machining, and thereby to facilitate these potential prospects, a fundamental understanding of the underlying material deformation and removal mechanism is desirable. Several relatively simple mechanisms will be discussed in this paper, supported by scanning electron and optical micrographs. Each has favourable operating conditions under which it will become dominant and therefore appropriate for the given set of conditions. In general, no process can be described by only one of these mechanisms and it is likely that several will co-exist on the same specimen.

2 Material Removal Mechanisms

2.1 Ploughing (Plastic Flow)

In the *ploughing mode*, material is plastically pushed aside by a tool rake face and a groove (furrow) is left behind it. Where there is no or very little evident swarf produced, the "removed" (displaced) material may have been redistributed to the surrounding bulk material via plastic flow or by densification (compaction) in the vicinity of the groove. In some materials such as soda-lime glass, pile-up provides

evidence of plastic flow. In silicon, a very high dislocation density area is found around indentations and scratched grooves (Puttick, Shahid and Hosseini 1979, Puttick and Hosseini 1980) showing traces of plastic flow.

This material removal mechanism, observed in scribing and ruling experiments, is normally obtained when the tool penetration depth is very small (of the order of 1 μm). Since the cutting depth has to be kept low, this mechanism is inefficient for bulk material removal, but arises under finish pass conditions.

2.2 "Cutting" Mechanism

The word *cutting* is used here to describe the shearing of materials from the bulk to form a continuous chip generated by its sliding away from the bulk along the rake face of the tool. It is the very effective cutting mechanism employed for traditionally-used ductile materials.

In ductile metal-cutting, it is well-established that there is a shear plane passing through a point of stress concentration present in the metal being cut. As bulk metal material approaches this shear plane, it undergoes substantial simple shearing as it crosses the thin shear zone (Shaw 1984). The chip then proceeds up the rake face of the tool and a rough surface develops on the back of the chip as the result of inhomogeneous strain involved in the process. To explain these observations, an ideal cutting process model, called the *card model*, was suggested by Piispanen (1969) which depicted the material cut as a deck of cards inclined to the free surface at an angle corresponding to that of the the shear plane. Although it does not explain chip curl or predict shear angle and contact length, it does model the chip-forming process.

Such chips in soda lime glass are shown in Fig. 1 with continuous and distinct shearing marks on the under side. This type of continuous-chip, called an "arc chip" or a "loose-arc chip", is similar to that produced when cutting conventionally ductile materials such as copper, mild steel and aluminium.

The chip cross-section is explained by the shear-angle associated with the central region of the chip being larger than that at the edge. For the chip to be continuous, part of it (centre) is under compressive stress whilst other parts (edges) are under tensile stress, the magnitude of these stresses depending on the cutting depth. When the cutting depth is small, the tensile stresses are not strong enough to initiate cracks so the chip edges present only small irregularity (as shown in Fig. 1(a)). Upon increasing the cutting depth, cracks start to propagate from both edges towards centre of the chip to form the "*centipede*" type chips (as shown in Fig. 1(b)). If the tensile stress is strong enough, cracks developed from both edges may meet at the centre to form a partial ring crack. In some cases, the edge detaches from the main chip body to form a long and thin ribbon.

In general, for brittle materials, the removal rate associated with this mechanism is much bigger than with *ploughing*. It was found in this study that the *cutting* mechanism can be obtained with a near-zero (small positive or small negative) rake angle tool. Use of a large negative rake angle, with resultant large compressive stress ahead of the tool, may suppress this type of material removal mechanism. For given conditions, there is a critical negative rake angle for the tool below which dominant *cutting* will be replaced by a dominating *ploughing* mechanism.

2.3 Delamination

Delamination is a process whereby ribbon-like swarf detaches from the workpiece bulk in the wake of the tool, and was proposed by Puttick et al (1989) as a ductile mode material removal mechanism (they observed this in machining soda-lime glass and fused silica (Spectrosil)). The *delamination* process is caused by the longitudinal compressive residual stress left by the tool immediately following its passage.

As shown in Fig. 2(a), the detachment of the swarf occurs by shearing fracture along a path within the stressed region parallel to the machined surface followed by bending or buckling caused by the compressive residual stresses on the end of the free segment. These "*hoops*" of swarf may be formed and extended to form a long ribbon which, in some cases, can be millimetres long (Fig. 2(b)).

In general, the observed *delamination* process in this study can be categorised into two types, namely *side-delamination* and *centre-delamination* . In the former case the material is forced to "pile-up" alongside the groove due to plastic-flow and then to detach itself from the bulk. In the case of *centre-delamination* , a layer of stressed material peels off from the surface after the tool has passed and the pattern of the groove is preserved on the curved swarf. Since different parts of the groove are subject to various degrees of compressive residual stress and the buckling is accompanied by large tensile stresses in the outer layers of the ribbon, the inhomogeneous stress field can cause the detached ribbon to split into several secondary ribbons curved in different directions according to the local stress field. From the viewpoint of dimensional control requirements, it does not appear favourable to remove material by *central-delamination* due to its lacking control of thickness of the detached ribbon and thence of the cut depth.

2.4 Brittle Fracture

By increasing the normal load on a tool or indenter, sub-surface lateral cracking might extend to the surface during unloading, causing part of material to be chipped off. These lateral cracks, as suggested by Hagans, Swain and Field (1979), are nucleated by shear deformation and driven by residual stress after the load is removed.

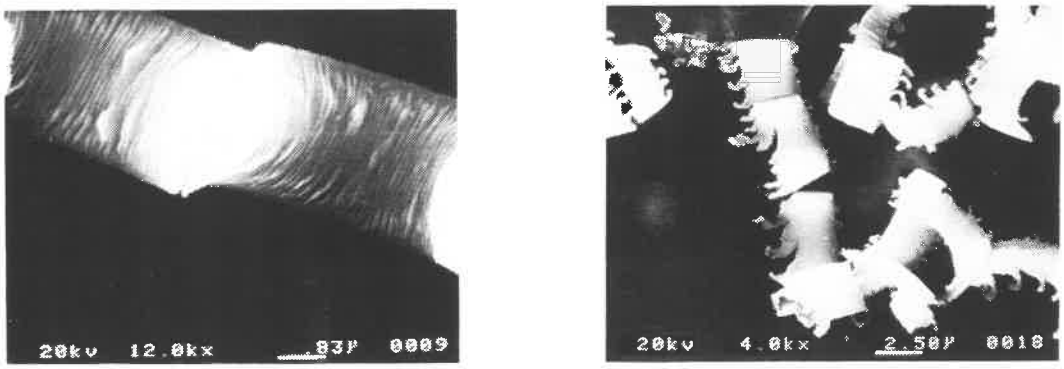
In the case of ruling and scribing, the tensile stress introduced by the friction force between the tool and workpiece can cause sub-surface lateral cracks to propagate along the cutting direction. As a result, it has also been found in ruling and scribing experiments on brittle materials that large chips (debris), over ten times larger than the groove size in some cases, could be produced by brittle fracture. Lawn and Wilshaw (1975) suggested that this mechanism is responsible for material removal during machining processes such as grit blasting or grinding.

Owing to the uncertainty of the crack directions and their extent, this material removal mechanism cannot offer the required dimensional accuracy in precision machining, although it can remove material very effectively. Also, the surface generated by using this mechanism is normally very rough with large scale sub-surface cracks (lateral and median cracks) penetrating deep ($\approx 20\,\mu m$) into the material.

Reference

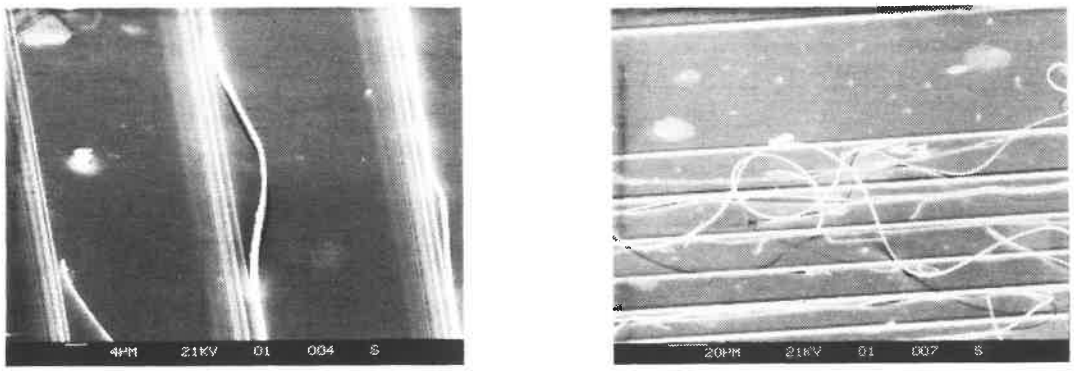
Bifano, T.G. Dow, T.A. & Scattergood, R.O.	1988	*Ductile-regime Grinding of Brittle Materials,* Proc. Conf. on Ultra-Precision in Manufacturing Engineering, Aachen, Germany
Bryan, J.B. & Carter, D.L.	1985	Lawrence Livermore Natl. Lab., Rep. No.: UCRL-92682, Berkeley, California
Chao, C.L. & Gee, A.E.	1989	IPES5/ASPE Meeting, 18-22 Sept., Monterey CA, Abstracts, p259
Hagan, J.T.; Swain, M.V. & Field, J.E.	1979	*NBS Special Publ. 562*, Ed. by B.J. Hockey and R.W. Rice, p15, U.S. Government Print. Office, Washington D.C.
Lawn, B.R. & Wilshaw, T.R.	1975	*J. Mat. Sci.*, Vol.10, p1049
Piispanen, R.	1969	*Ann. C.I.R.P.*, Vol.XVII, p561

Puttick, K.E. & Hosseini, M.M.	1980	*J. Phys. D: Appl. Phys.*, Vol.13, p875
Puttick, K.E. Shahid, M.A. & Hosseini, M.M.	1979	*J. Phys. D: Appl. Phys.*, Vol.12, p195
Puttick, K.E. et al	1989	*Proc. Roy Soc. A.* 426, p19
Shaw, M.C.	1984	*Metal Cutting Principles*, Oxford University Press, London



(a) (b)

Fig. 1 (a) small cutting depth: the tensile stresses are not strong enough to initiate cracks so the chip edges present only small irregularity. (b) increasing the cutting depth, cracks start to propagate from both edges towards centre of the chip to form the "*centipede*" type chips.



(a) (b)

Fig. 2 (a) Example of the smaller swarf from soda lime glass which has both ends still attached to the groove edge and has its middle section arched up to form a "*hoop*" shape (b) large curly swarf were observed after ruling soda lime glass.

Fixed-Abrasive Grinding of CVD SiC to Optical Quality

Thomas Bifano and Yuan Yi
Aerospace and Mechanical Engineering
Boston University
110 Cummington St.
Boston, MA 02215
(617) 353-5619

Current Research at Boston University on Ultraprecision Machining of SiC

Chemically vapor-deposited silicon carbide (CVD SiC) is a leading candidate material for large, space-based high energy mirror applications. Because of its high energy specific stiffness, 99.9999 percent purity, and structural stability, CVD SiC is also being used in less exotic applications, such as in the production of lens molds and replicated optics. Polished CVD SiC is an excellent reflector, particularly for 11-12.5 μm wavelengths [3]. Occasionally, ground CVD substrates have been coated with silicon in a vapor deposition process [4]. The silicon layer can then be *diamond-turned* to produce required figure and finish in a single operation.

This paper describes high-precision diamond grinding of as-deposited polycrystalline CVD SiC to optical quality. It will be shown that the Si cladding added for diamond turning purposes may not be necessary, since CVD SiC can be ground in the ductile regime to surface finishes of 10 $\mathring{A}$ R_a, with no subsurface damage. One of the problems encountered in fixed-abrasive grinding of CVD SiC is dulling of the grinding wheel, and this phenomenon will be described later.

Preliminary experiments were performed on an air-bearing based ultraprecision grinding apparatus. Both infeed and crossfeed were controlled in real time with a bandwidth of 1 KHz and a resolution of 5 nanometers using capacitance gage feedback and piezoelectric actuaton in a micro-computer-based control algorithm. The coolant was de-ionized water. The CVD SiC test samples were ground under the following conditions:

Sample Area	$5mm \times 25mm$
Infeed per pass	$1\mu m$, $100nm$
Cross-feed	$1.25mm/min$, $7.7mm/min$
Coolant	De-ionized water
Wheel Speed	1500 RPM

The ground CVD SiC surfaces were examined using a scanning electron microscope (SEM). SEM photographs of the ground CVD SiC reveal that the surface is damage-free, as illustrated in Figure 1. In these photomicrographs, the only visible features are cutting marks spaced about 600 nm apart, corresponding to the cross-feed of the

workpiece per revolution of the grinding wheel. No subsurface damage is visible. In the surface grinding operation used to prepare these samples, grinding damage extended more than 40 μm deep in this sample.

For the grinding geometry used in these experiments (wheel speed much greater than cross-feed speed), the specific grinding energy, μ, is given by:

$$\mu = \frac{F_T \cdot V_w}{MRR}$$

where:

F_T is tangential grinding force (obtained from the torque of motor), N
V_w is grinding wheel peripheral speed, m/s
MRR is material removal rate, $mm^3/mm \cdot s$

For up to $4mm^3$ of material removal, μ is plotted in Figure 2. μ has an initial rise for the first mm^3 of material removal, is essentially steady for the next $2mm^3$ of material removal, and rises dramatically after $3mm^3$ of material removal. Eventually, increased forces will result in a more damaged surface as forces between the grinding wheel and the workpiece increase. Currently under study is a technique for continuous re-dressing of the resin bond diamond grinding wheel to continually expose new diamonds at a controlled rate to prevent excessive grinding force while maintaining machine accuracy.

The ground CVD SiC surfaces were also examined using stylus profilometry. Figure 3 shows a 500 μm trace along the surface. Over that length the sample had a roughness of 11 $\AA$ R_a.

From preliminary grinding experiments, it has been determined that for CVD SiC, mirror finishes can be acheived by ductile-regime grinding. Moreover, diamond grinding of SiC has an advantage over glass in the production of large optics because of its greater ductility and because the CVD surface is reflective without post-machining coating or polishing.

From this preliminary data, the potential for fixed abrasive grinding of CVD SiC optical components without silicon cladding, diamond turning, or post-polishing has been established.

Further experiments are being carried out at Oak Ridge National Lab on the Productivity Validation Test Bed (PVTB), using a ultraprecision grinding apparatus. In these experiments, aspheric optics are to be generated on CVD SiC by fixed-abrasive ductile-regime grinding.

References

[1] Bifano, T. G., Ductile-Regime Grinding of Brittle Materials, PhD Dissertation, North Carolina State University, 1988.

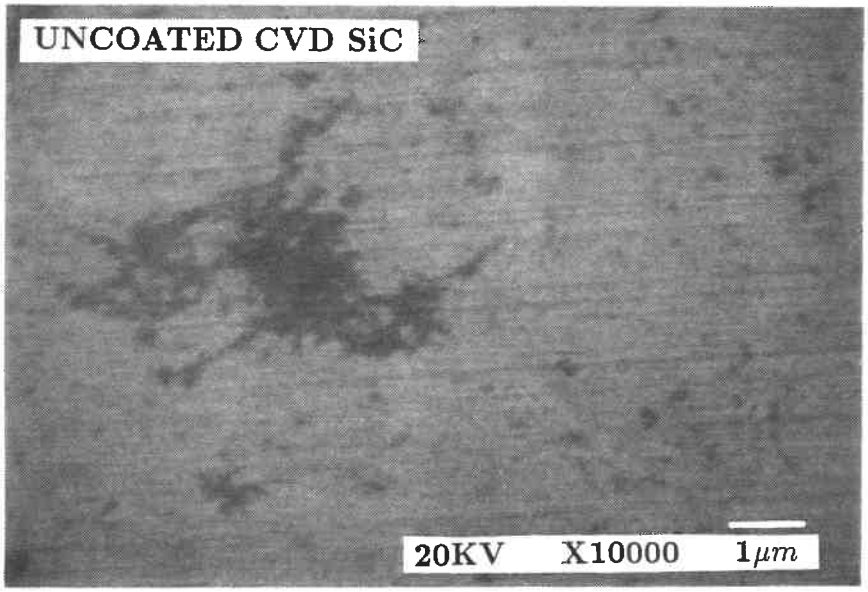


Figure 1: SEM photographs of the ground CVD SiC

[2] Blackley, W. S. and Scattergood, R. O., Ductile-Regime Maching Model for Diamond Turning of Brittle Materials, **Precision Engineering Journal**, V.13, No.2., April, 1991.

[3] Pickering, M. A., SiC-An Emerging Material, **Proceedings of Fifth Industrial briefing**, Oak Ridge National Laboratory, pp. 321-344, Nov 9-10, 1989.

[4] Casstevens, SiC-An Emerging Material, **Proceedings of Fifth Industrial briefing**, Oak Ridge National Laboratory, pp. 345-365, Nov 9-10, 1989.

[5] Bifano, T. G., and Fawcett, S. C., Specific Grinding Energy as an In-Process Control Variable for Ductile-Regime Grinding, **Precision Engineering Journal**, to be published in 1991.

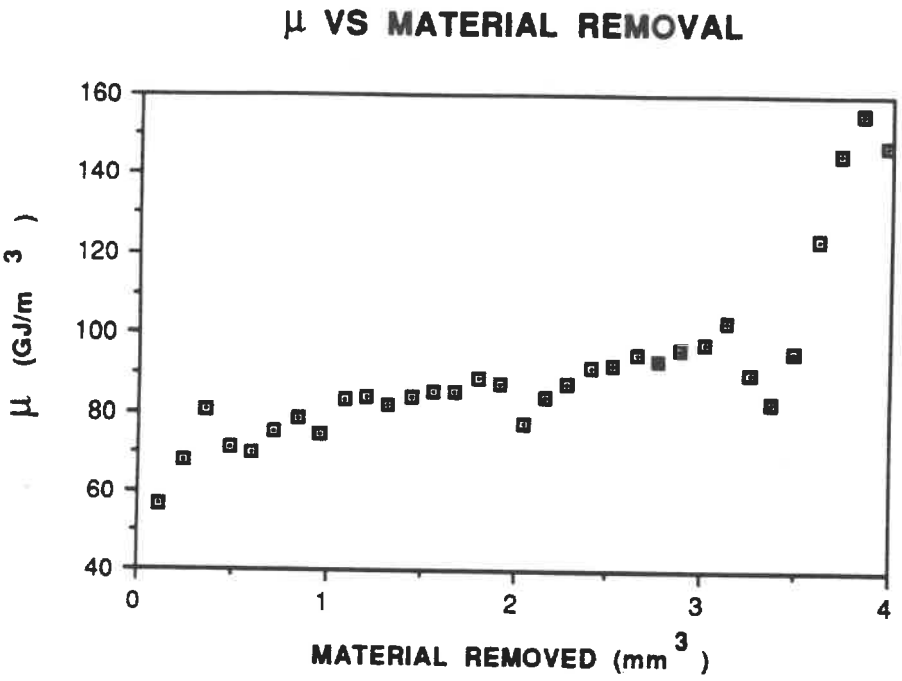


Figure 2: Specific Grinding Energy for Grinding of CVD SiC

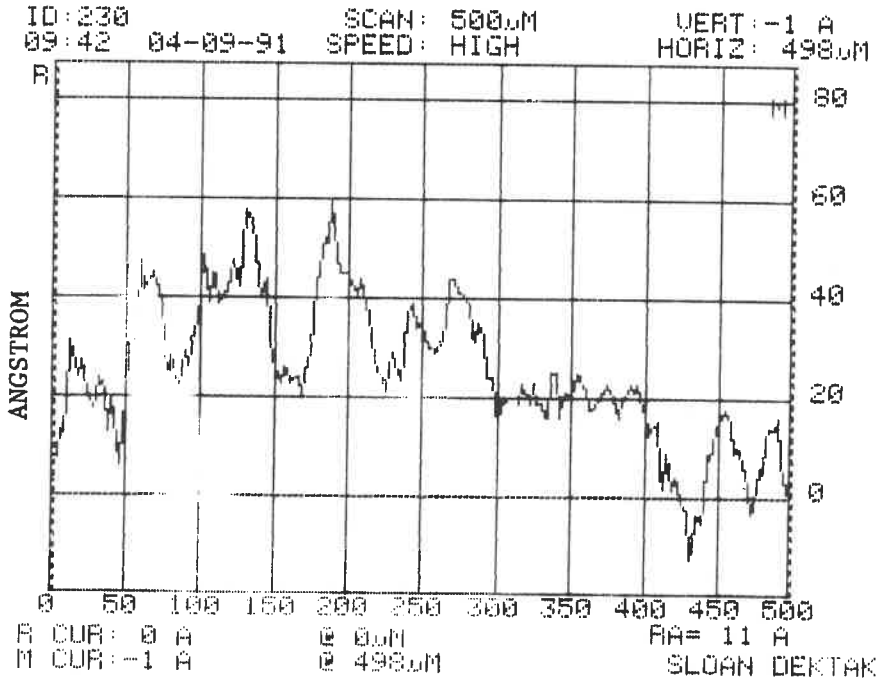


Figure 3: Profile of as-ground CVD SiC surface.

Wheel Speed Effects in Precision Contour Grinding

Steven C. Fawcett[†]
Thomas A. Dow

Precision Engineering Center
North Carolina State University
Raleigh, NC 27695

Introduction

A geometric computer model of a precision grinding operation was developed to calculate the surface features generated during contour grinding with a radiused wheel. The model indicates that small changes in the wheel speed can have a dramatic effect on the surface finish. Analysis of ground surfaces reveals uniform surface profiles and easily distinguished features which could only be produced by a constant wheel speed. This occurs for an air-bearing, air-turbine grinding spindle which has limited torque and is driven under open-loop control. The effects of the relative speeds are analyzed and an energy based **"phase locking"** mechanism is proposed that can provide feedback to the grinding spindle from the material removal operation.

Grinding Model

Computer simulations [1] can predict the influence of the relative spacing of discrete cuts formed during a precision grinding operation. Figure 1 illustrates two variations, the In Phase condition where wheel speeds equal integer multiples of the workpiece rotational speeds ($N_w = n\ N_p$) and the Out of Phase condition. where wheel speeds are integer multiples plus 1/2 of the part speed ($N_w = (n + 1/2)\ N_p$). Two-dimensional models cannot include the effects of these interactions.

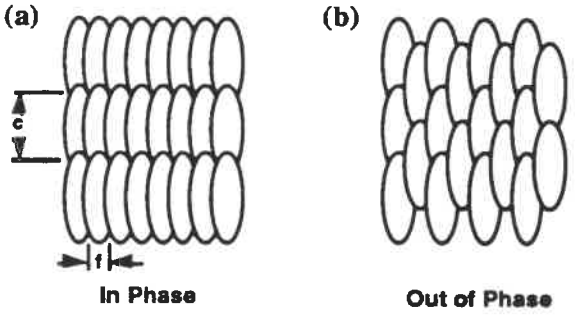


Figure 1 Diagram of the cuts made by the high spots on the grinding wheel.

By calculating the surface roughness over a range of grinding wheel speeds, the effect of the phase variation becomes apparent. Figure 2 depicts the theoretical roughness calculated from the computer simulations of the grinding interaction as a function of wheel speed. The roughness exhibits a cyclic response to relatively small changes in grinding wheel speed. The period of the response is equal to rotational speed of the workpiece (90 RPM).

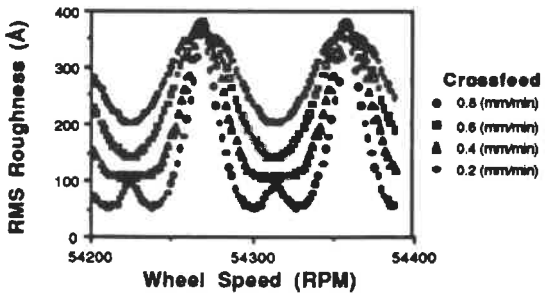


Figure 2 Plot of the theoretical roughness versus grinding wheel speed.

[†] Currently at NASA, Marshall Space Flight Center, Alabama

The theoretical roughness shows maximas at the In Phase condition. This graph also shows how the minimum points vary as a function of the crossfeed. For this particular geometry, crossfeeds above 0.4 mm/min have the minimum at the Out of Phase condition. The simulation of the Out of Phase surface appears to have twice the number of cuts as the same distance in the In Phase surface.

There is a noticeable transition in the surface roughness data of Figure 2 for crossfeeds less than 0.4 mm/min. There is no longer a single minimum location. The minimum now occurs at two wheel speeds ($N_w \approx (n \pm 1/3)\ N_p$). This transition is designated as the Branching Point (BP). Because at these points the nose radius of the grinding wheel is extremely large relative to the crossfeed rate, the surface appears to have three times the number of cuts as the same distance on the In Phase surface.

The effect of the phase interaction on the chip dimension is shown in Figure 3. The phase shifts have an inverse effect on the chip dimensions as they did on the surface finish.

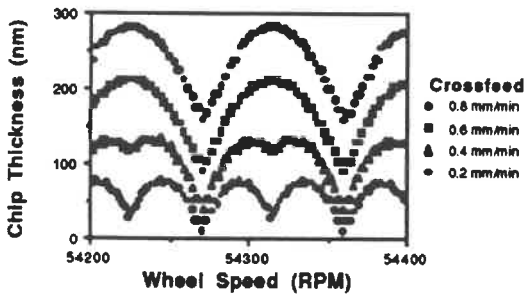


Figure 3 Plot of the theoretical chip dimensions versus grinding wheel speed.

The branching point (BP) has major significance when examined from the chip thickness perspective. When the data exhibits the branching effect, the variation in the chip thickness is decreased. Since spindle speeds typically cannot be controlled very accurately, a smaller deviation in the chip dimensions over a range of speeds would be desirable.

Branching Point Parameter

The six dominant grinding variables are the part speed (N_p), wheel speed (N_w), nose radius (A), wheel radius (R), crossfeed rate (F) and part diameter (D). A systematic evaluation of the influence of each of these variables was undertaken to accurately predict the onset of branching. This investigation produced the combinations of variables for a parameter denoted as the Branching Point (K_{BP}) [2].

$$K_{BP} = \left(\frac{A\, D^2\, N_p^2}{R\, F^2 \sqrt{N_w}} \right) \qquad (1)$$

For example, a particular geometry should exhibit branching if the parameters give a value of the branching point (K_{BP}) greater than 5000. If this point is less than 4000, the geometry will not exhibit the branching effect. These results are summarized in Table 1..

	Minimum Surface Roughness	Maximum Chip Thickness
$K_{BP} > 5{,}000$	$N_w \approx (n \pm 1/3)\ Np$	$N_w \approx (n \pm 1/3)\ Np$
$K_{BP} \approx 4{,}500$	Transition	Transition
$K_{BP} < 4{,}000$	$N_w \approx (n + 1/2)\ Np$	$N_w \approx (n + 1/2)\ Np$

Table 1 Branching point summary.

Grinding Experiments

Evaluation of precision ground ceramics indicated roughness values near the simulated minimum. Figure 4 shows the measured roughness for ground CVD SiC and the limits calculated from the simulations. The measured roughness follows the minimum calculated for the case where the wheel speed is $N_w = (n\pm1/3)\ N_p$ with an offset of about 7 nm RMS. The fact that the measured values follow the same trend as the calculated minimum suggests that the grinding spindle maintains a constant speed.

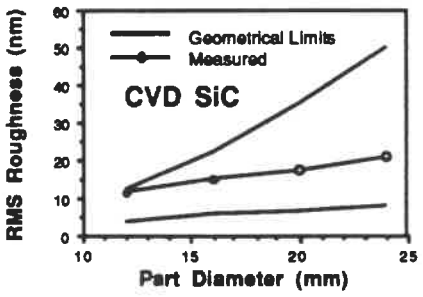


Figure 4 Plot of the measured and simulated surface finish for the CVD SiC sample.

Other ground surfaces corroborated the relationship between the measured surface roughness and the calculated minimum as illustrated in Figure 4. Because the minimum roughness corresponds to a specific ratio of the wheel speed to the part speed, this ratio must remain constant to achieve these values. The lack of feedback control for the grinding spindle suggests that the grinding process itself must be providing the feedback mechanism.

Material removal by brittle fracture requires less specific energy than removal by plastic deformation [3]. In the contour grinding experiments, the majority of the material removal proceeds by brittle fracture even though a small portion of the chip produced over the finished surface is produced by plastic deformation. Because less energy input is required for the case of brittle fracture, the grinding process would theoretically try to maximize the volume of material removed by fracture and minimize the volume removed by plastic deformation.

Grinding experiments were performed to determine if the wheel speed would "lock" into the speed predicted by the analytical derivation. During the grinding operation, the wheel speed and the part speed were constantly measured. One experiment was performed during the grinding of HIP SiC with a calculated branching point (K_{BP}) of 5,200. This value is above the threshold where the data starts to exhibit branching so the minimum roughness should occur where $N_w = (n\pm1/3)\ N_p$. The wheel speed was measured during the experiment and the results are plotted in Figure 5. The theoretical ratio (n) is calculated as:

$$(n + 1/3) = \frac{N_w}{N_p} \approx \frac{55{,}400}{89.1} \Rightarrow n = 621 \qquad (2)$$

Also shown in Figure 5 are other values of the ratio for n±1/3 for reference. From this the theoretical wheel speed where the maximum chip thickness occurs is calculated.

$$N_w = (n + 1/3)\ N_p = 55{,}370.0\ \text{RPM} \qquad (3)$$

The average measured speed (55,370.4 RPM) is almost identical to the theoretical calculation.

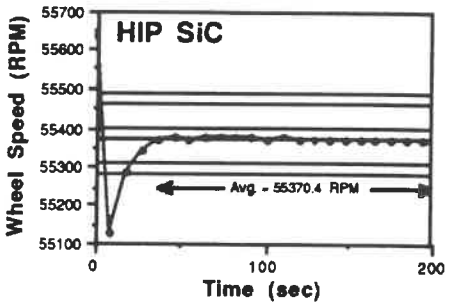


Figure 5 Wheel speed measured during the grinding of HIP SiC.

The surface of the ground sample was also examined to determine if the phase effects caused by the constant spindle speed could be seen on the surface. Figure 6 shows a Nomarski photograph of the ground sample. The alternating dark and light regions correspond to high and low regions formed by the discrete cuts in the cutting direction. The fact that these regions are aligned over many crossfeeds verifies that the wheel speed did remain constant. The angle that these lines form with the crossfeed marks is approximately at 17 degrees. This corresponds to a wheel speed to part speed ratio of N_w = (n+1/3) N_p and also corroborates the theory.

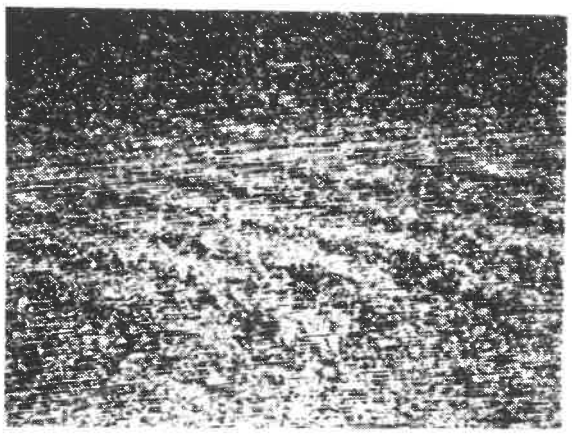


Figure 6 Nomarski micrograph of ground HIP SiC sample (130X).

Conclusions

The wheel speed to part speed ratio is an important variable in the model of contour grinding. Speed changes as small as 0.1% can have a dramatic effect on the magnitude of the surface roughness changing it by a factor from 2X to 8X depending on the feed rate. Experimental verification of the magnitude of the surface finish and the uniformity over the surface of the part indicated that a mechanism must exist to keep the wheel speed at a constant ratio with the part speed. The geometric model provided a relationship between the speed ratio and the chip dimensions that led to an energy-based explanation of the consistency of the wheel speed. Measurements of the wheel and part speed confirmed this hypothesis.

References

1. Fawcett, S.C. and T.A. Dow. "Precision Contour Grinding of Brittle Materials", **Precision Engineering**. (Accepted for Publication), 1991.

2. Fawcett, S.C. "Development and Implementation of a Grinding Technique for Precision Finishing of Brittle Materials" Ph.D. Dissertation, N. C. State University, 1991.

3. Bifano, T.G. and S.C. Fawcett. "Specific Grinding Energy as an In-Process Control Variable for Ductile-Regime Grinding", **Precision Engineering**. (Accepted for Publication), 1991.

Experimental machining of glass to optical finishes in a closed loop configuration

C. Gary Blough, Erwin G. Loewen, and Duncan T. Moore
The Institute of Optics
University of Rochester
Rochester, New York 14627

To diamond turn glass to a specular finish, the depth of cut must be controlled in the nanometer domain. This requires significant modifications to a standard diamond turning machine. Normal lead screw mechanisms cannot control infeed to the required level, even with interferometric control. In addition most diamond turning machines lack the rigidity and the environment for long term thermal stability needed. The stiffness of the structural loop must be increased because glass is a hard and brittle material, leading to relatively high cutting forces. Nanometer stability normally requires ambient temperature control to the millidegree level, which is impractical as well as expensive.

Fortunately a relatively simple approach exists that, for experimental purposes at least, provides both freedom from thermal disturbances as well as tight position control. It consists of a non-contact proximity sensor system that effectively connects the tool holder to the spindle face, and controls a piezo-electric pusher that determines the location of the diamond tool to 20 nm. The resulting toolholder, based on the Cranfield design, is shown in Figure 1.

Previous work[1] indicated that to machine glass in a ductile manner required depths of cut in the range of 0.1-1.0 μm. Therefore, we used workpieces that were polished 50 mm diameter plates, held parallel to better

than 5 arc seconds. However, this still left over 1 μm of wedge with respect to the tool's path. Establishing first contact with certainty is essential to efficient operation. A sensitive accelerometer located on the tool itself gave a usable signal within 50 nm of first contact.

To test the fidelity of the closed loop system required a hard brittle metal, beautifully represented by germanium, (which has approximately the same hardness as glass). The tool could be programmed in a complex path and the grating like shape accurately measured. The result was confirmed to a few nanometers, from which it was concluded that the structural loop stiffness and accuracy should be entirely sufficient to machine glass.

Using steep negative rake tools(-45°- -60°) many glass samples were machined to surface roughness values as low as 15 Å Ra. However, a common feature was that the material actually removed was only a fraction of its expected value. Plunge cutting tests in aluminum revealed that tool wear was much too small to account for the discrepancy.

Compliance of the system seemed the most likely explanation, yet the signal from the feedback control loop was extremely stable during machining. Plunge cutting the tool in germanium, both before and after, pointed to the solution. It revealed that cutting just a few meters of glass degraded the tool edge sharpness enough so that it could no longer cut germanium, but eventually led to brittle fracture when the load was sufficiently increased. Therefore, the previously mentioned machining of glass looks to have been more burnishing than cutting.

The use of germanium as a reference material has proven very useful even though glass machines quite differently. Future work will concentrate on

identifying the cause of the astonishingly fast edge degradation. The origin is believed to be chemical in nature because of the absence of edge chipping or evidence of elevated temperatures on the glass surface. We plan to investigate the source of such an effect. The role of coolant or inert atmospheres around the tool will also be tested.

Acknowledgements

The authors wish to thank the New York Center for Advanced Optical Technology and the Army Research Office for support of this research.

In addition, the authors gratefully acknowledge the contributions from IBM, Rank-Pnuemo, and Cranfield Unit for Precision Engineering whose help has made this research possible.

References

1. M. Schinker and W. Doll, "Grundlegende untersuchungen zur spanenden bearbeitbarkeit von anorganischen glasern mit einzahnwerkzeugen" , Fraunhofer-Institut fur Werkstoffmechanik, Freiburg, FRG.

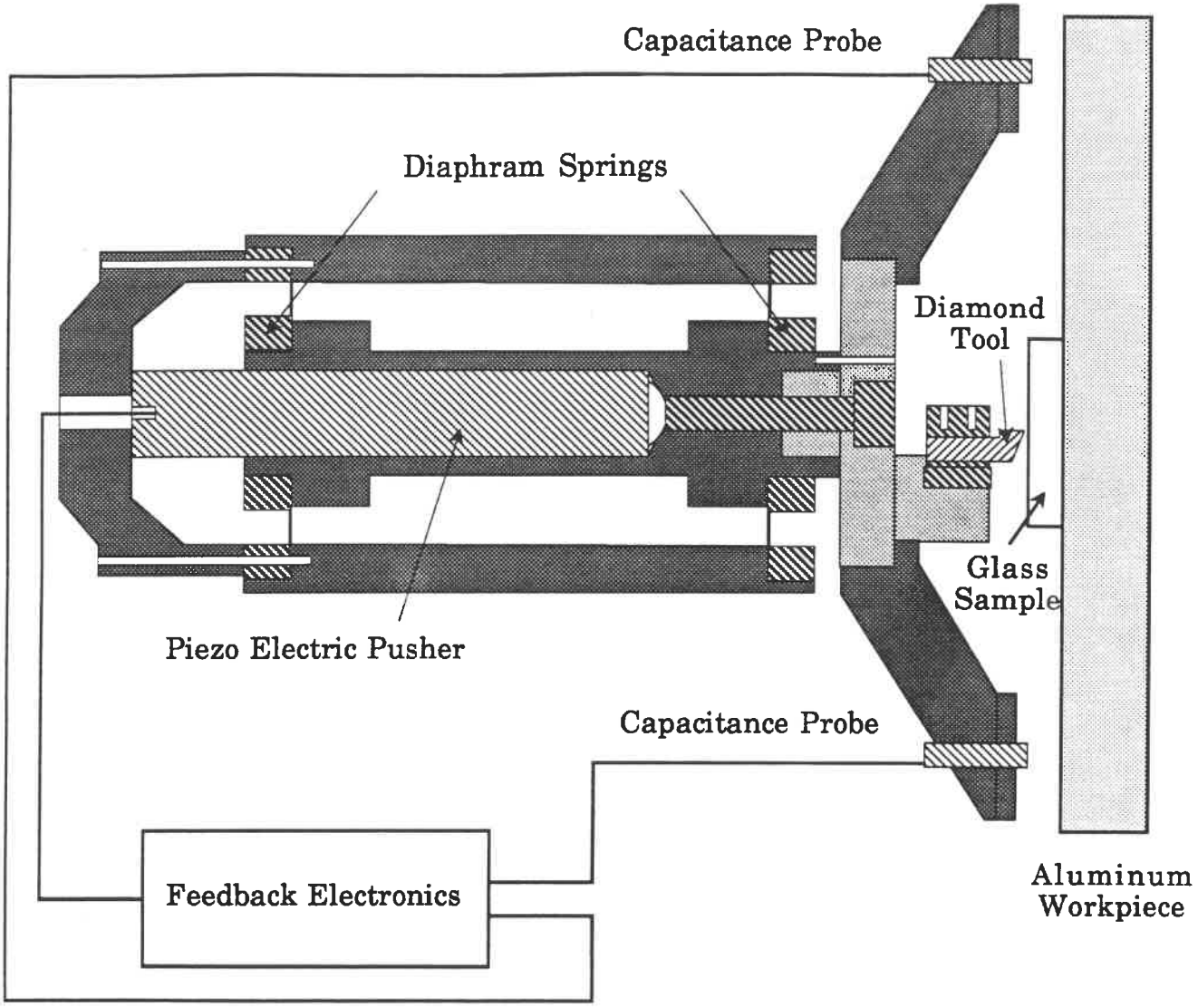


Figure 1. Nano-Positioning Tool Holder

Grinding of Optical Glass with Loose Polycrystalline Synthetic Diamond Abrasives

M. J. Cumbo*, A. Lindquist, S. B. Ng, T. Rich, Y. Sabharwal and S. D. Jacobs**

The Institute of Optics
University of Rochester
Rochester, NY 14627

EXTENDED ABSTRACT:

The purpose of this study is to characterize the grinding properties of borosilicate optical glass using loose polycrystalline synthetic diamond abrasives and small-scale plano production equipment. Polycrystalline synthetic diamond has been advertised in trade journals [1,2] to promote high rates of glass removal with low levels of subsurface damage (SSD). Experiments in our laboratory have demonstrated that, for a given removal rate, grinding of borosilicate glass with polycrystalline diamond abrasives results in a significant reduction in both SSD depth and surface roughness relative to grinding with several common monocrystalline diamond abrasives.

The loose abrasives used in this study are commercially available monocrystalline diamonds (natural, synthetic metal bond and synthetic resin bond) and synthetic polycrystalline diamonds. [3] The natural and polycrystalline diamonds are generally intended for use in loose abrasive grinding while the metal and resin bond diamonds are usually incorporated into bound abrasive grinding wheels, ring tools or pellets.

Two sizes (3 μm and 1 μm) of each of the four diamond types were evaluated as loose abrasives suspended in deionized water (1% by weight) in the grinding of BK-7 glass with a fused silica lap. A Strasbaugh Model 6DE single spindle machine, typical of small-scale equipment used by many companies for precision plano glass finishing, was used for all grinding experiments.

---

* Bausch and Lomb Inc.
** Laboratory for Laser Energetics, University of Rochester.

The process parameters, fixed throughout this study, were as follows. The diamond/water slurry was dispensed at 20 ml/min. The pressure applied to the fused silica lap, which was oriented on top of the BK-7 work, was 1.23×10^{-3} kgf/mm^2 (1.76 psi). The spindle and eccentric speeds were 39 rpm and 47 rpm, respectively.

To ensure identical initial surface conditions, the BK-7 work was preconditioned by grinding with 5 μm alumina in deionized water (10% by weight) prior to each diamond grinding experiment. Using the above parameters for 60 minutes, this resulted in the removal of about 75 μm of glass from the work.

For each of the eight diamond abrasives, we evaluated the glass removal rate, surface roughness, SSD depth and mode of glass removal. The results are summarized in Table 1. Surface roughness was measured with a Talystep Model 112/1037 profilometer. The radius of the stylus was 2 μm and the load was 1 mgf. SSD depth was determined using the Subsurface Tribometer method developed previously in our laboratory. [4]

Under identical process conditions using similarly sized abrasive particles, polycrystalline diamond abrasives provide comparable glass removal rates with monocrystalline diamond abrasives, but with superior surface quality. The surface roughness values and depth of SSD of the polycrystalline diamond ground BK-7 are the lowest of the four diamond types for both abrasives sizes.

In loose abrasive processes, the depth of SSD is typically of the order of the abrasive size used to grind the surface. In the case of the 3 μm abrasives, this relationship holds for the monocrystalline diamonds, but not for the 3 μm polycrystalline diamond. The polycrystalline diamond particles break apart as they wear, thereby exposing new sharp edges. This results in a depth of SSD characteristic of a smaller monocrystalline abrasive without the corresponding penalty in the rate of glass removal. In the case of the 1 μm diamond abrasives, the typical relationship between abrasive size and SSD is violated. This is due to the fact that an insufficient amount of glass was removed during each of the 1 μm grinding experiments to clear the SSD caused by the preceding 5 μm alumina grinding operation. With the relatively low removal rates obtained, substantially longer grinding times would be necessary with 1 μm diamonds to properly determine the depth of SSD.

Consistent with the above results, after 60 minutes of grinding with 1 μm polycrystalline diamond, the BK-7 surface reflectance was highly specular, indicating ductile removal. None of the other abrasives produced such a smooth surface.

The mode of glass removal (brittle fracture or ductile shearing) is of interest because of the potential utility of ductile grinding as an intermediate step between the grinding and polishing stages of conventional glass finishing processes. With proper control of induced surface stress, ductile grinding may provide a means of reducing the total polishing time required in the finishing of precision optics. As shown in this study, the use of micron-sized synthetic polycrystalline diamond abrasive provides a means of achieving ductile removal of BK-7 glass with reasonable removal rates using conventional grinding equipment.

ACKNOWLEDGEMENTS:

This study was supported by the New York State Center for Advanced Optical Technology (NYSCAOT) and the Center for Optics Manufacturing (COM).

REFERENCES:

[1] Bailey, N. F., "Explosion-Synthesized Polycrystalline Diamond Powder: Its Manufacture, Characteristics and Uses," Proceedings: Diamond's in the 80's, a Technical Symposium, (1980), Industrial Diamond Association of America.

[2] Bailey, N. F., "Unique Values of Polycrystalline Diamond Powder in Polishing Superhard Materials", Intersociety Symposium on Machining of Advanced Ceramic Materials and Components, (1988), American Society of Mechanical Engineers.

[3] Available from Amplex Corporation, Bloomfield, CT 06002.

[4] A. Lindquist, S. D. Jacobs and A. Feltz, "Surface Preparation Technique for Rapid Measurement of Subsurface Damage Depth," Science of Optical Finishing, Optical Society of America, 1990 Technical Digest Series, Vol. 9.

TABLE 1

Summary of BK-7 Grinding Results With Diamond Abrasives

Diamond Type	Average Removal Rate[a] (μm/min)	Peak-to-Valley Surface Roughness (μm)[b]	RMS Surface Roughness (μm)[b]	SSD Depth (μm)[c]	Glass Removal Mode
Errors	±0.002	±25%	±25%	±0.5	
3 μm Diamonds					
Natural	0.571	1.167	0.186	3.7	brittle
Resin Bond	0.253	1.189	0.168	2.7	brittle
Metal Bond	0.623	1.267	0.177	2.9	brittle
Poly	0.497	0.821	0.141	1.4	brittle
1 μm Diamonds					
Natural	0.085	0.402	0.073	2.8	brittle
Resin Bond	0.046	0.386	0.062	2.7	brittle
Metal Bond	0.104	0.260	0.049		
Poly	0.040	0.122	0.020	1.9	ductile

a) Average of 60 minutes of grinding time, evaluated at two minute intervals.
b) Average of three 1 mm scans.
c) Average of five sites.

SHEAR MODE GRINDING OF BK7 GLASS AND EVALUATING GROUND SURFACE

J. Takeda, K. Imai, H. Hashimoto
Kanagawa Institute of Technology
1030, Shimo-ogino, Atsugi, Kanagawa 243-02, Japan

In order to clarify pragmatic aspects of shear mode grinding, grinding processes and characteristics of ground surface were investigated.

First grinding experiments were carried out, using #400*, #1000M, #3000M, #5000M, and #1000R** wheels and BK7 specimen (50x3x4mm).

Next, ground surfaces were characterized by a variation in the rate of etching and a super microindentation. Results obtained were summarized briefly as follows:

1) Corresponding to the number of infeed, in shear mode grinding with metal bond wheels, first, the normal grinding force increases linearly and later reaches a stable state which shows a constant grinding force and no further residual stock removal. However, in #1000R grinding that shows shear mode surface, the normal grinding force shows a high increase and no stable state can be found.
2) Further, in the spark-out process, the normal grinding force of #3000M shows an exponential decrease to around 3N, whereas that of #1000R does not show the rapid decrease.
3) Comparison of a #5000M ground surface and a polished surface was made by vickers super microindentation test, (load: 10mN-100mN, loading speed: 1mN/sec, holding time: 10sec).
   As for the results, both a ground surface and a polished surface are harder in the very near surface. It is found that the ground surface is softer than the polished surface for the first 0.1 μm.
4) The etching rate of a shear mode surface produced by #3000M grinding is higher than that of a polished surface for the first 0.2 μm. However, in the range deeper than 0.2 μm, the etching rate was found to be the same.
5) Tips of abrasive grains were blunted after truing, and in the early state of the shear mode grinding, the surface roughness of the grinding wheel #3000M enlarges from 0.1 μm to 0.3 μm. Thereafter, that change is hardly observed in the experimental range (until 1000 times of infeed). During this experiment the necessity for dressing was not observed.

* # 400M: Metal bond wheel impregnated with 400 mesh diamond.
** #1000R: Resinoid bond wheel impregnated with 1000 mesh diamond.

A MIRROR-FINISH SLICING TECHNOLOGY WITH ELECTROLYTIC DRESSING FOR HARD AND BRITTLE MATERIALS

Torahiko Kanda and Masashige Mitsuhashi
Functional Devices Research Laboratories, NEC Corporation,
1-1 Miyazaki 4-chome, Miyamae-ku, Kawasaki, Kanagawa, 216 JAPAN

1. Introduction

The optical waveguide switch/modulator used with single mode optical fiber has recently shown dramatic promise for expanding optical communications systems. Fig.1 illustrates typical light coupling from optical fiber to optical waveguide. Efficient light coupling requires mirror-finish endfaces to avoid optical loss. As almost all optical devices are made of hard and brittle materials, such as $LiNbO_3$ and SiO_2, their endfaces are usually lapped and polished after conventional slicing with a thin blade. This process is time-consuming, however, and both time and cost might be saved with a new slicing technology which could produce mirror-finish endfaces without the need for lapping and polishing.

In general, optical device endfaces are considered sufficiently smooth, when their surface roughness Rmax has been reduced to approximately 0.05μm. In order to obtain such a surface by slicing alone, the mesh size of a micro-grain blade would have to be at least #4000 (4μm). Metal bond blades are well-suited to such purposes, because they tend to maintain even micro-grain blade surfaces. In actual practice, however, dressing is extremely difficult. An electrolytic dressing method for mirror-finish surface grinding with a metal bond wheel has been reported[1]. Sufficient protrusion of micro-grains can be realized by bond metal electrolysis effect. This paper proposes a mirror-finish slicing technology, which uses a thin metal bond blade(#4000) to which an electrolytic dressing is applied, for use with optical devices. An adequate electric current was also decided upon by experiments.

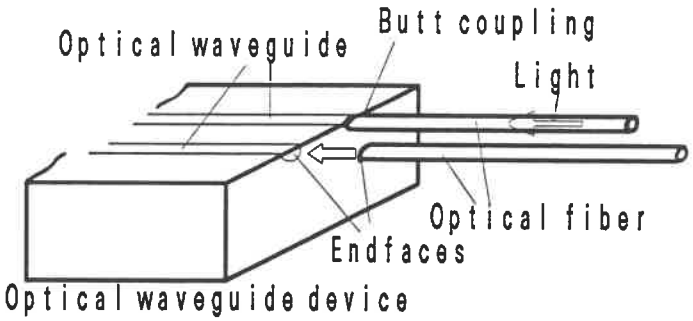


Fig.1 Light coupling between optical fiber and waveguide

2. Electrolytic Dressing with Newly Developed Minus Electrode

The configuration for slicing with a thin blade is different from that for surface grinding[2]. This can be attributed to the following: (1)low rigidity for thin blade, (2)relatively large grinding forces, due to large contact surface between blade and workpiece, (3)and some right angle error for the workpiece feed direction to the spindle axis.

These features cause an elastic deformation of the thin blade during the slicing process. The workpiece endface, sliced mainly by the blade peripheral working face, is also considered to be slightly ground off (rubbed) by the blade side face. The sliced endface quality may be markedly influenced by the blade side face grinding ability. Since the blade side faces remove only a small amount from the workpieces, the self-dressing effect is hardly expected. Therefore, it is important for the blade side face to retain its grinding ability appropriately, undergoing only a little electrolytic dressing. Regarding these points, the shape of the minus electrode, which can accomplish an appropriate electrolytic dressing on the blade side faces as well as the blade peripheral face, was investigated during several previous experiments.

Figs.2 and 3 show the newly developed minus electrode and the experimental set-up. The gap distance between the minus electrode and the blade side face is set much larger than gap d between the minus electrode and the blade peripheral face. Thus, a small electric current, and then a little electrolytic dressing is fed to the blade side faces.

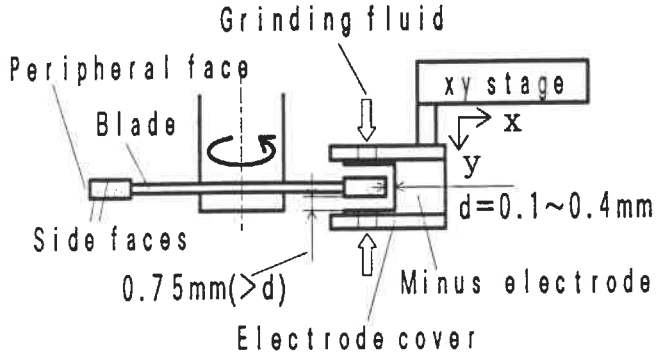


Fig.2 Minus electrode cross section view

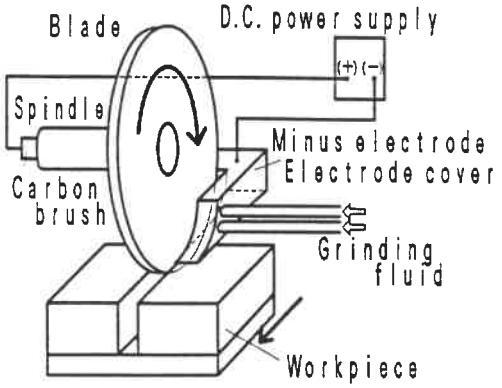


Fig.3 Experimental set-up

3. Experimental Method and Results

3-1 Experimental method

Table 1 shows the main slicing conditions. A bronze bond blade is mounted on the air-spindle, as shown in Fig.3. A carbon brush set, which is connected to the 60V D.C. power supply, is attached to the other end of the air-spindle. Electric current was controlled, so as to be at a set value by changing gap d, as shown in Fig.2. Surface roughness Rmax values for the sliced endface, using the proposed slicing technology, a conventional slicing without electrolytic dressing, and so on, were measured by TALYSTEP, TALYSURF, and SEM, as required. Also, the optical loss, caused on the sliced endface, was evaluated.

Table 1 Slicing conditions

Slicing Blade : SD4000 Bronze Bond 75 (76X0.5X40)
Workpiece : Lithium Niobate (Width:10mm)
Depth of Cut : 4mm, Feed rate : 3mm/min (Down cut)
Spindle Revolution : 15000rpm (V:3600m/min)
Grinding Fluid : Water Soluble Type (3ℓ/min)
Electrolytic Dressing : Applied Voltage DC 60V

3-2 Results and discussion

3-2-1 Side face dressing effect

Figs.4(a) and (b) are SEM photographs of the blade side face after a couple of slicing passes, without and with electrolytic dressing. In Fig.4(a), it is observed that grains are embedded in the bond metal, whereas sufficient protrusion of grains is found in spite of micro-grain, in Fig.4(b).

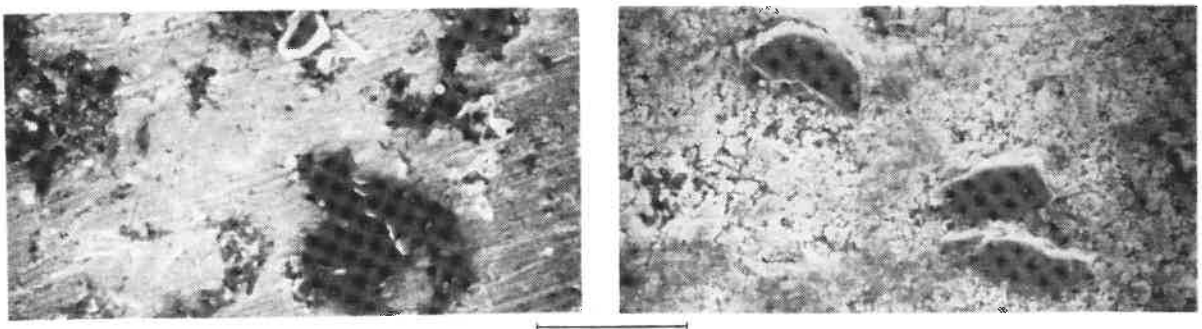


Fig.4 SEM views of the blade side face (a)without electrolytic dressing, (b)with electrolytic dressing

Fig.5 shows the change in workpiece surface roughness Rmax, in slicing $LiNbO_3$ substrate with three different electrolytic dressing conditions. Although these test results had some dispersion, the average

surface roughness was nearly 50 % smaller, when electrolytic dressing was applied, than when it was not applied. Without electrolytic dressing, many micron size chippings resulted from brittle fracture are found on the entire sliced endface, as seen in Fig.6(a). When electrolytic dressing is fed only to the blade peripheral face by using usual minus electrode, micron size chippings still remained, [see Fig.6(b)]. Chipping is presumably caused by wheel loading and/or rubbing by the bond metal on the blade side face.

On the other hand, when electrolytic dressing is applied, using the developed minus electrode, no chippings appeared, as shown in Fig.6(c). Instead, grooves made with plastic flow were observed on the entire sliced endface. The average surface roughness Rmax value was made still smaller, and a mirror-finish endface (Rmax 0.06μm) could be realized. These indicate that the grinding ability of the blade side face plays an important role, and that it is appropriately maintained by electrolytic dressing using the developed minus electrode. Similar results were obtained in slicing SiO_2 optical fibers which are fixed to an appropriate glass substrate.

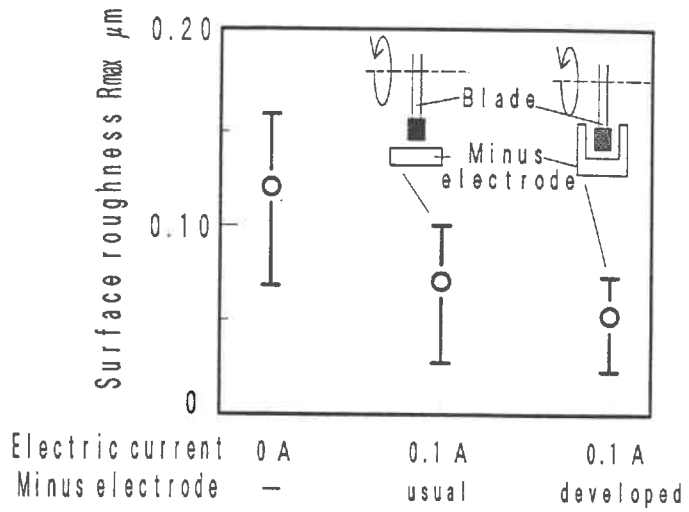
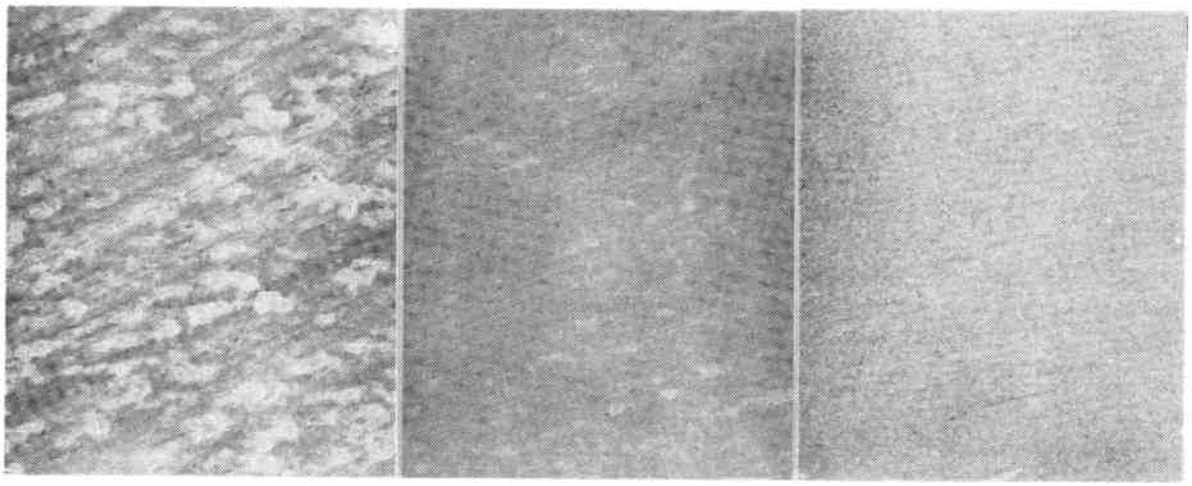


Fig.5 Decrease in surface roughness by electrolytic dressing



(a) Without electrolytic dressing (b) Usual minus electrode 10μm (c) Developed minus electrode

Fig.6 SEM photographs of sliced $LiNbO_3$ endface

3-2-2 Adequate electric current

In the proposed mirror-finish slicing technology, more electrolytic dressing improves the blade grinding ability, and then decreases the sliced surface roughness. It is inferred, however, that excess electrolytic dressing would cause rapid blade wear. To find the adequate electric current, experiments were carried out at four different electric currents.

Fig.7 shows the change in surface roughness Rmax with the increase in slicing passes. Before the first slicing pass in each electric current, the blade was trued and dressed by 0.2A electric current for 3 minutes. This pre-dressing was conducted to make sure the blade surface had a uniform and appropriate grinding ability. In Fig.7, the surface roughness with 0.2A electric current is the smallest, then comes 0.4A,

0.1A, and finally 0A. It was also found that surface roughness Rmax is low and constant in the case of 0.2A, whereas Rmax increases as the slicing passes number increases, in the cases of 0.1A and 0A.

Electric currents 0A and 0.1A are insufficient for maintaining the blade side face grinding ability appropriately. Nevertheless, surface roughness does not increase immediately after the first slicing pass. It is considered that wheel loading and/or rubbing by the bond metal on the blade side face gradually took place. Then, surface roughness was deteriorated with the increase in number of slicing passes. This indicates that the blade side face needed a little electrolytic dressing.

When 0.4A electric current was applied, several pitches were found on the sliced endface. These pitches resulted from rapid blade wear, owing to excess electrolytic dressing. From these results, it was experimentally derived that adequate electric current was 0.2A.

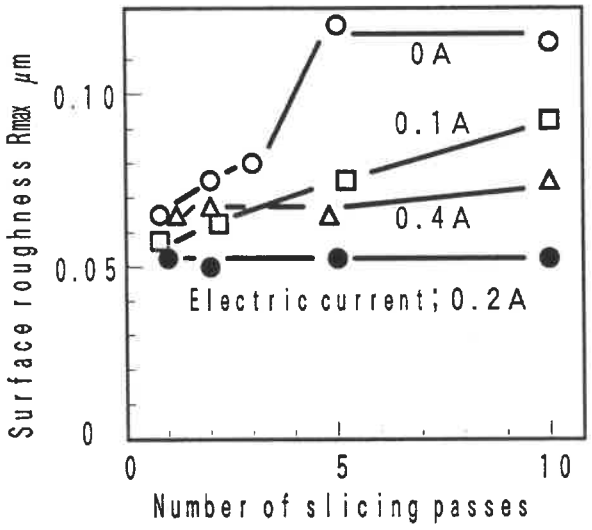


Fig.7 Surface roughness with number of slicing passes

3-2-3 Optical light transmission characteristics

In order to assess the sliced optical device endface by mirror-finish slicing technology, optical insertion loss, including optical loss caused on the sliced endface, was measured. By using mirror-finish slicing technology, insertion loss was greatly improved, in comparison with that achieved using conventional slicing. The insertion loss, caused on mirror-finish endfaces by this technology, was a little inferior to that achieved by lapping and polishing processes. However, excess loss was less than 0.5dB for $LiNbO_3$ waveguide endface and was less than 0.1dB for optical fiber endface. These small losses caused no problem in practical use. Therefore, lapping and polishing processes can be eliminated. Machining time for endfaces by mirror-finish slicing technology was less than one tenth of that when using lapping and polishing processes.

4. Conclusion

This study has proposed a mirror-finish slicing technology for hard and brittle optical devices, which uses a thin metal bond blade applied with electrolytic dressing. A new shape minus electrode was developed and an adequate electric current was decided upon by the experiments. The optical device endface, sliced by this technology, is so smooth and has such reasonable properties in terms of optical light transmission that lapping and polishing processes can be eliminated.

Acknowledgements

The authors wish to thank Dr. T. Nakagawa, Professor at Tokyo University, Dr. S. Esho, H. Fukuchi, and F. Yamamoto for their valuable advice. The authors also thank Mr. A. Toyohara for carrying out part of the experimental work.

References

1. H. Ohmori, T. Nakagawa : Mirror Surface Grinding of Silicon Wafer with Electrolytic Inprocess Dressing, Annals of the CIRP, 39-1(1990) 329
2. S. Matsui, K. Syōji, H. Teramoto : Studies on Outer-blade Slicing(1st Report) Journal of JSPE, 53-07(1987) 1051 [in Japanese]

A LONG RANGE STM/AFM

Claude Lavigne, Jimmie A. Miller and Robert J. Hocken
Precision Engineering Laboratory
University of North Carolina at Charlotte
Charlotte, NC 28223

INTRODUCTION

Most high resolution microscopes such as Scanning Tunneling Microscopes (STM) or Atomic Force Microscopes (AFM) are limited in image range to a few tens of microns [1,2,3]. The purpose of our study was to build a STM/AFM set allowing us to work on large surfaces up to 150 x 150 microns, with the possibility of atomic resolution, and use this device to characterize engineering surfaces. This new tool extends the range of applications for those microscopes, and offers the possibility of comparative studies with other instruments such as optical interferometers, profilometers, etc.

CONCEPT

The basic concept was to build a long range STM using a commercial (X,Y,Z) piezo transducer macro-stage [4,5]. The principle is to use different flexure arrangements to amplify the motion of a piezoelectric driving element to obtain nanometer resolution. A flexure stage is made of a series of pivoted lever arms which amplify the motion by the ratio of the lever arms lengths. The sample to be studied is fixed upon the Z axis (Fig. 1). This stage gives us a first orthogonal coordinate system (X,Y,Z) allowing three dimensional motion in a volume of 150 x 150 x 18 microns, with a resolution of a few nanometers. This stage is used first to get an image at medium resolution for large surfaces, then in high resolution mode to move the chosen area under the tip. In the case of the STM, the sample is swept (X,Y) in a raster fashion to produce a two dimensional topography of the surface. An applied potential between the tip and sample produces a current which is monitored for feedback height control of the sample (Z). By keeping the tunneling current constant, rough surfaces can be characterized. The head of the STM is made of a tip holder mounted on a scanning tube. Here we used a commercial product. We chose this type of head in order to get an additional orthogonal coordinate system (x,y,z). Thus we have now an STM with 6 degrees of freedom. The first three (X,Y,Z) providing motion of the sample, the last three (x,y,z) motion of the tip.

CONTROL

We designed our own electronics interface in order to control the tunneling current between the tip and the sample. Information obtained from the feedback determines the displacement of the sample (Z). The bandwidth of the analog feedback is roughly 15 kHz. The microscope was calibrated with a gold plated diffraction grating and checked with an uncalibrated Optical Linewidth Standard

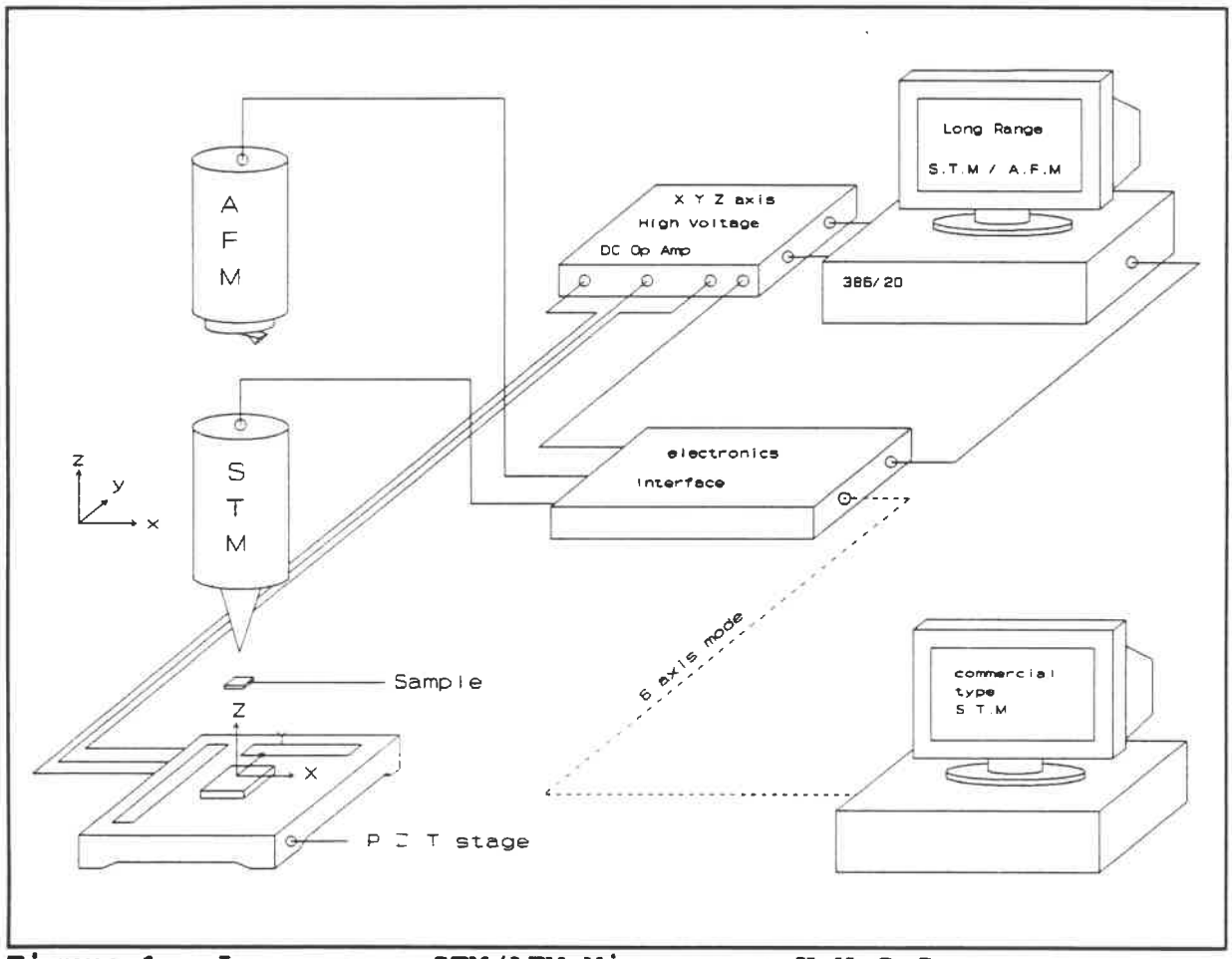


Figure 1. Long-range STM/AFM Microscope U.N.C.C.

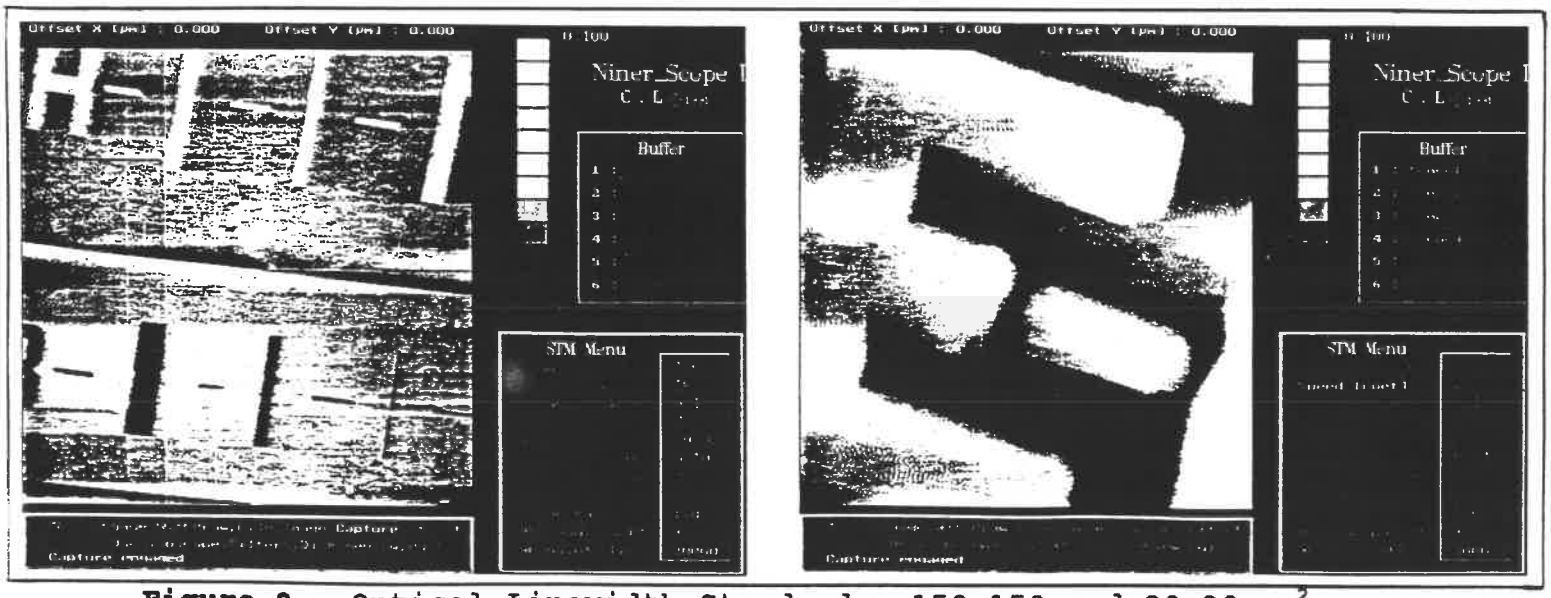


Figure 2. Optical Linewidth Standard: 150x150 and 30x30 μm^2.

(Fig. 2) [6]. Among the various applications of this system is scanning large surfaces at medium resolution, then "zooming in" to

get selected images at high resolution. We can also scan the same area using successively one of the two coordinates systems (X,Y,Z) or (x,y,z). We were careful during the writing of the software to provide different ways to scan one sample. Indeed one purpose of this study was to develop a system to test different types of microscopes, so the software was written to obtain images using different scanning modes. We can scan in X or Y, as well as in reverse for each axis in order to study the consequences of the asymmetries on an STM tip, or the effect of an AFM's cantilever. AFMs are particularly sensitive to scanning direction and speed. The system can also be used for servo characterization by imaging the same area scanning in different directions and comparing the results.

AFM DESIGN

Once the STM had been used for several surface characterizations, we built an AFM using the same base. The design allowed the alternate use of the two types of microscopes on the macro-stage without excessive manipulation. Force measurements were performed by monitoring the deflection of a cantilever, resulting from the inter-atomic forces between atoms on the cantilever tip and atoms on the sample's surface. In our case the deflection of the cantilever is held constant by servoing the sample height. The AFM signal was obtained from the exponential variation of the tunneling current between a Pt-Ir tip and a gold film on the backside of the lever [7]. Our earliest cantilevers were made of a gold coated aluminum foil cut by hand. Tips for these cantilevers were made by gluing a very short STM tip (200 microns) to the end of the lever. The vertical spring constant and mechanical resonant frequency of the Aluminum lever were calculated to be 30 N/m and 10 kHz. Using this first design we were able to get several images with a resolution of 15 nanometers in a long-range scanning mode. To achieve a better resolution we needed to reduce the spring constant and maintain a high resonant frequency. To do this we decided to use a previously demonstrated fabrication technique for constructing thin film cantilevers on silicon substrates [8]. This method allows us to build cantilevers with high resonant frequencies (greater than 10 kHz) and a low force constant in order

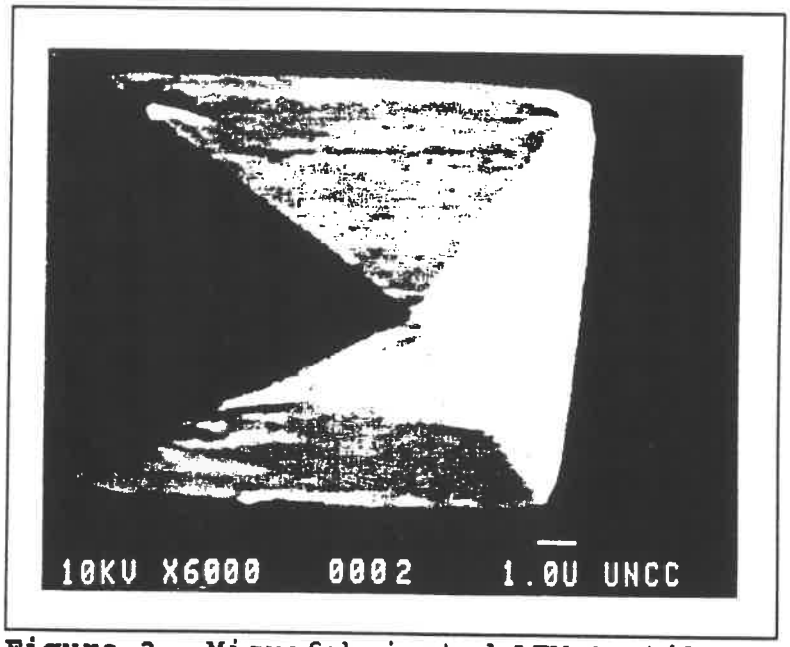


Figure 3. Microfabricated AFM Cantilever Tip

to measure forces of less than 10^{-8} Newton. A tip fabricated in this manner is shown in Fig. 3.

CONCLUSION

A new 6 axis STM/AFM microscope has been designed, fabricated, and successfully tested by imaging several samples. The first purpose was to develop a long range microscope with an alternative to the piezoelectric tube for the scanning mode. We choose this linear stage to avoid the errors induced by the bending of the PZT tube for large displacements. In addition, having an AFM upon the same site provides the possibility of working with virtually all types of samples, conductive or not. Ongoing improvements in AFM tip fabrication should provide us with the possibility to resolve atomic features. Early results on selected surfaces will be presented.

ACKNOWLEDGEMENTS

The authors would like to thank H. Estrada and B. Boudreau for their help on the microfabrication of the AFM cantilevers, and R. Larrabee of NIST for his assistance with the photomask Standard. This work was supported by the National Science Foundation.

REFERENCES

[1] G. Binnig, C.F. Quate and Ch. Gerber, "Atomic force microscope," Phys. Rev. Lett. 56,930 (1986a).

[2] Y. Kuk and P.J Silverman. "Scanning tunneling microscope instrumentation," Rev. Sci. Instrum. 60(2), February 1989.

[3] Dror Sarid, "Scanning Force Microscopy", Oxford series on optical sciences; Oxford university press 1991.

[4] F.E. Scire and E.C Teague, " Piezodriven 50 micrometer range stage with subnanometer resolution", R.S.I 49 1735,1973.

[5] R.D. Young, "Micromeasurements with the piezo-flex Stage", WYE CREEK INSTRUMENTS, Frederick, MD 21701

[6] R. Larrabee, "Bright-Chromium Linewidth Standard, SRM 476, for Calibration of Optical Microscope Linewidth Measuring Systems.", N.I.S.T.

[7] Tatsuya Ishizaka, Yasushiro Sugawara, Kozo Kumagai and Seizo Morita, " Anomalous Force Dependence of AFM Corrugation Height of a Graphite Surface in Air", Japanese Journal of Applied Physics, Vol. 29, No. 7, July, 1990, pp. L 1196-L 1198.

[8] Albrecht, T. "Advances in atomic force microscopy and scanning tunneling microscopy." PhD Dissertation, 1989; Stanford university.

two laser diodes of wavelength 670nm and 780nm (the use of 2 wavelengths subsequently enables diffraction order numbers to be identified) coupled into a low birefringence single-mode fibre terminated by an achromatic collimating lens. The purpose of this fibre is (i) to act as a spatial filter by selectively transmitting the Gaussian TEM 00 mode and (ii) to provide a convenient means of producing spatially coincident wavefronts.

For both wavelengths, the Gaussian beam is then directed at normal incidence through a cylindrical lens onto the surface to form a line image at the beam waist (to restrict diffraction to one plane only). The diffraction pattern arising from the phase modulation of the incident wavefront at the surface is then focused onto a 2048 CCD detector. Finally, the subsequent data processing is carried out on a dedicated pc which also enables diffraction patterns, power spectra and surface parameters to be displayed.

3.0 TALYFINE PERFORMANCE

The results of various tests using the first production Talyfine instruments have been encouraging. The components measured were diamond-machined brass (Naval non-leaded), aluminium (6061 T6) and copper (OFHC) flats. They were initially measured (normal to the lay) on a Talystep fitted with a nominal 0.1μm tip radius chisel diamond stylus and data-logged at 1μm spatial intervals. The resulting data sets were then Fourier transformed and filtered in frequency space to achieve the same bandwidth as the Talyfine measurements (ie a selectable cut-off filter together with a 12.8μm Heaviside cut-on filter). From this data, the 'stylus-instrument' (Talystep) values for the power spectra together with the parameters R_q, Δ_q and λ_q were evaluated and compared with the equivalent Talyfine values.

Discussing these in turn, there is a very good agreement between the two sets of filtered power spectra, although it is clear when measuring such surfaces with the Talyfine, the distribution of power within the power spectrum does vary over the surface, the greatest variations being due to form errors.

Similarly, there is good agreement when comparing the R_q values, as shown in fig 2. It is noteworthy that from this data, no significant material dependence is apparent.

A similar comparison between the 2 sets of λ_q values for these surfaces shows noticeably more scatter. When measuring the surfaces at various locations with Talyfine it is clear that, whereas R_q is very stable, λ_q varies significantly (as does Δ_q).

An additional set of 'High-R_q' samples (up to 145nm) consisting of both gratings and diamond-machined cylinders were similarly measured (on a Form Talysurf) and their R_q's compared with those of the Talyfine. Such surfaces give rise to multi-order diffraction patterns, so that the derivation of surface power spectra are no longer possible. However, good agreement for the 2 sets of R_q values were obtained, although with more scatter than the low R_q results.

As a separate exercise, to verify the principle that when using diffraction as a measurement technique, sample rotation is irrelevant, a germanium flat was measured on a diamond machine

SURFACE CHARACTERISATION via OPTICAL DIFFRACTION

Daniel Mansfield
Rank Taylor Hobson Ltd
Leicester, UK

1.0 INTRODUCTION

This paper explores the application of diffraction as a non-contact means for measuring high quality surfaces such as those produced by single point diamond machine-tools. Through considering the surface as a reflective phase grating, the various relationships that exist between the statistical parameters of the surface topography and the far-field diffraction pattern are established. These relationships have been incorporated into a new measuring instrument designed primarily for the measurement of flats and cylinders, the Talyfine[1]. The performance of this instrument is compared against an established stylus-based surface texture measuring instrument, the Talystep[1].

Starting with a formulation of the relevant Kirchoff integral, the diffraction process is explored with particular emphasis on the various statistical metrological parameters that may be extracted from the diffraction pattern: R_q, Δ_q, λ_q (rms values for surface roughness, surface wavelength and surface gradient), together with the surface power spectra. Of particular import is the low-Rq regime, this being typified by diffraction patterns dominated by the zero order; within this regime, it is legitimate to directly filter the diffraction pattern in a manner analogous to multiplicatively filtering the power spectrum in Fourier space, apart from the additional proviso that the energy 'lost' is transferred to the zero order. These filters may be chosen to be directly equivalent to the industry standard metrology filters[2]. The ensuing benefit in the form of stability against point-spread function distortion due to surface waviness is discussed.

The Talyfine instrument in fact has two separate sets of algorithms; those concerned with diffraction patterns that are dominated by a single order and those that are not. These are discussed in turn.

In the paper, an overview of the Talyfine instrument is presented followed by a detailed analysis of the relationships existing between the far-field diffraction pattern and the surface topographical statistics. Finally, a comparative assessment of Talyfine is given with the stylus based instrument, the Talystep[(1)].

2.0 TALYFINE INSTRUMENT OVERVIEW

A Talyfine schematic is shown in Fig.(1). Referring to this, the delivery system is based upon

tool; no dependence on spindle rotation speed was observed.

4.0 CONCLUSION

In the paper, an overview of the Talyfine instrument is presented together with a detailed analysis of those relationships that exist between the far-field diffraction pattern and the surface topography. The virtue of using diffraction to characterise low Rq surfaces has been demonstrated by the agreement between the Talyfine instrument and the stylus-based Talystep. In particular, the strong advantages achieved through directly filtering the diffraction pattern are shown. In addition, measurements of Rq on a rotating surface have confirmed that, unlike a stylus instrument, there is no distinction between static and dynamic measurements.

5.0 ACKNOWLEDGEMENTS

The author is grateful to the directors of Rank Taylor Hobson Ltd for permission to publish this paper.

6.0 REFERENCES

(1) Rank Taylor Hobson Ltd, Troon Ind. Est., Leicester, UK.

(2) International Standard ISO 327H-1975(E) 'Instruments for the Measurement of Surface Roughness by the Profile Method - Contact (Stylus) Instruments of Consecutive Profile Transformation - Contact Profile Meters System M'.

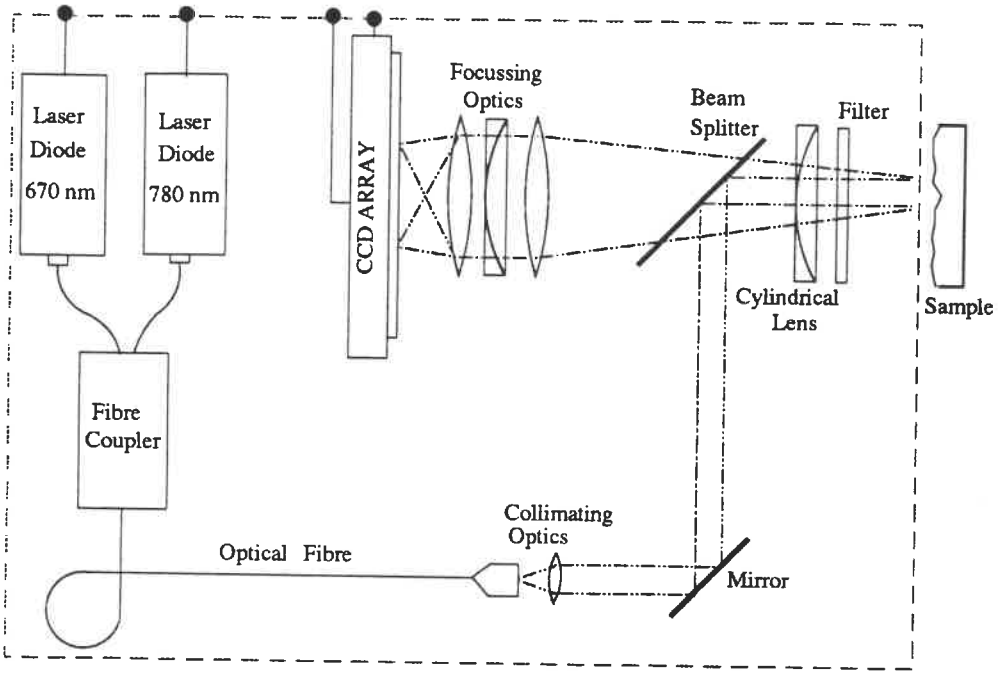


Figure 1 Schematic of Talyfine

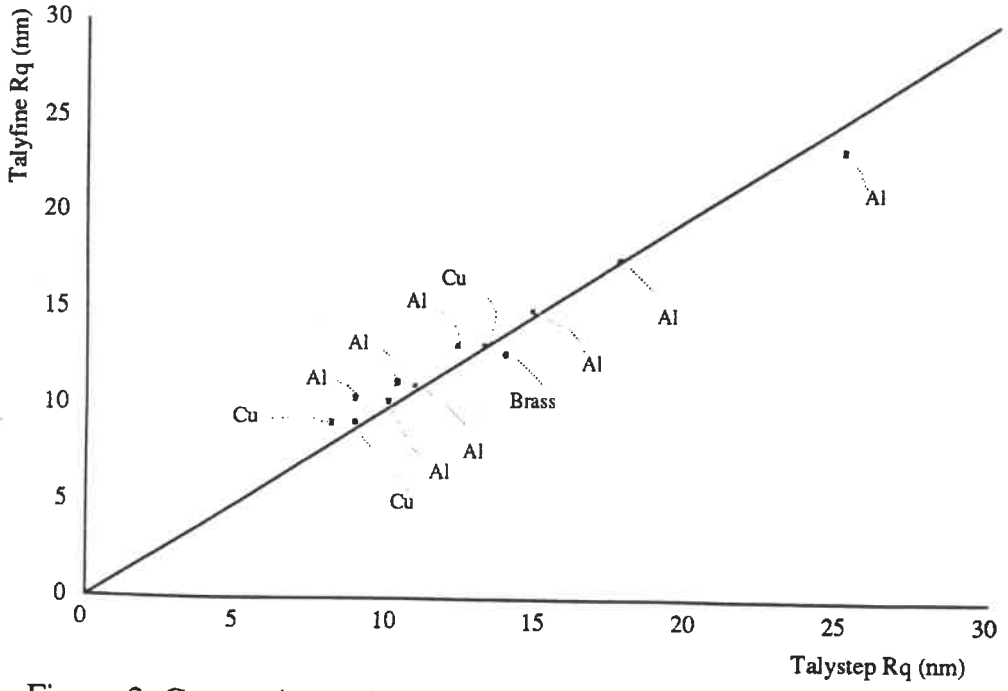


Figure 2 Comparison of Talyfine with Talystep: rms Roughness Rq

THE MEASUREMENT VARIATION OF SURFACE PROFILE PARAMETERS : A VIEWPOINT FROM STANDARDISATION

T Y Lin, P J Sullivan, and K J Stout
School of Manufacturing and Mechanical Engineering
University of Birmingham, UK

Kai-Li Ko
Centre for Measurement Standards
Industrial Technology Research Institute
Hsinchu, Taiwan, R.O.C.

Introduction

The measurement variation of surface profile parameters has become an important topic among surface metrologists and standardisation researchers. To date, stylus-based instruments have been the most commonly used devices to obtain profile information. Their popularity is predominantly due to their simplicity of operation and the embodiment of many of their operating principles in international standards. When using a profile measuring system, variation is generally found in the parameters acquired at different sites from a surface. This parameter variation is due to many causes (see below) including the operation of the instrument itself. The lack of understanding of the causes of parameter variation has limited the abilities of stylus instruments in obtaining a more complete representation of surface topography.

When using a digital profile measurement system, several operating factors must be selected before a measurement can be made. Such factors include sampling conditions i.e. the spacing of ordinates in a sample, the assessment length, and the size of the sample. In the study of sampling variation of a surface parameter, it is of great value to develop expressions of the sampling distribution as a function of these operating factors. Most of the analytical studies of sampling variation do not consider the effects of these three operating factors completely. For instance sampling interval is a predominant factor in obtaining a discrete sample from a continuous surface. Unfortunately, the influence of this critical factor has not been included in the theoretical expressions of sampling variation from a correlated or an uncorrelated stochastic process.

Causes of measurement variation when using a stylus-based digital profile instrument

When taking measurements at different sites (or indeed the "same" site) from a surface using a stylus-based digital profile measuring system, variations are found in the parameters acquired. The major causes of this variation may be classified as follows:

(A) **Variation in the geometry of the real machined surfaces.**
There are many possible causes for variation in the geometry of a machined surface. Some causes are due to manufacturing conditions (influenced by tool wear

and damage, and the presence of cutting debris). The variation may be described by the surface geometric integrity (homogeneity, isotropy, stationarity, and normality, etc.).

(B) **Variation in the digitisation of real machined surfaces.**
There are many factors which determine a digital measurement systems ability to obtain information about the real continuous surface. These factors may be classified into (i) inherent limitations of the measurement process (e.g. quantisation error) and (ii) selectable measurement conditions (e.g. choice of sampling conditions).

(C) **Inherent variation in the characterisation parameter.**
Some estimates of surface parameters are more unstable due to their parameter integrity. For example, the variation of Rt (peak-to-valley) values is larger than that of Ra (average) estimates.

Although parameter variation is dependent upon all of the above contingencies, standardisation attempts to limit the variation by proposing guidelines. Thus, embodied within many international standards are guidelines for stylus geometry and load, filter type and cut-off, and profile orientation. However, at present there are no guidelines to enable standardisation of sample spacing, and sample size. In this study we will show how these digital operational variables effect the parameter variation obtained when all other contingencies remain constant.

Scope and results of this investigation

In this paper, the variation of sampled geometric features from sinusoidal or stochastic surfaces under specified sampling conditions is discussed, based on the assumptions that the effects of machining variability and instrument limitation are insignificant.

The discussed profile parameter, representing the geometric features of the sampled profile, is mainly based on the sample mean square due to its importance in the characterisation of a profile. The square root of the sample mean square, Rq, is recognised as one of the primary characterisation parameters with meaningful statistical properties.

In dealing with the variation extent of sample mean square, analytical expressions considering the key factor of sampling interval are derived for the AR(1) and AR(2) surfaces. Here an AR(1) or AR(2) surface is the surface from which the sampled profiles can be represented by the corresponding AR(1) or AR(2) model, which can characterize many machined surfaces. In addition the theoretical results with infinite discrete ordinates from sinusoidal surfaces are presented.

The influences of sample length, sample spacing, and sample size on the measurement variations from sinusoidal and random surfaces are investigated based on the developed analytical expressions. Some results of this study are shown in following figures.

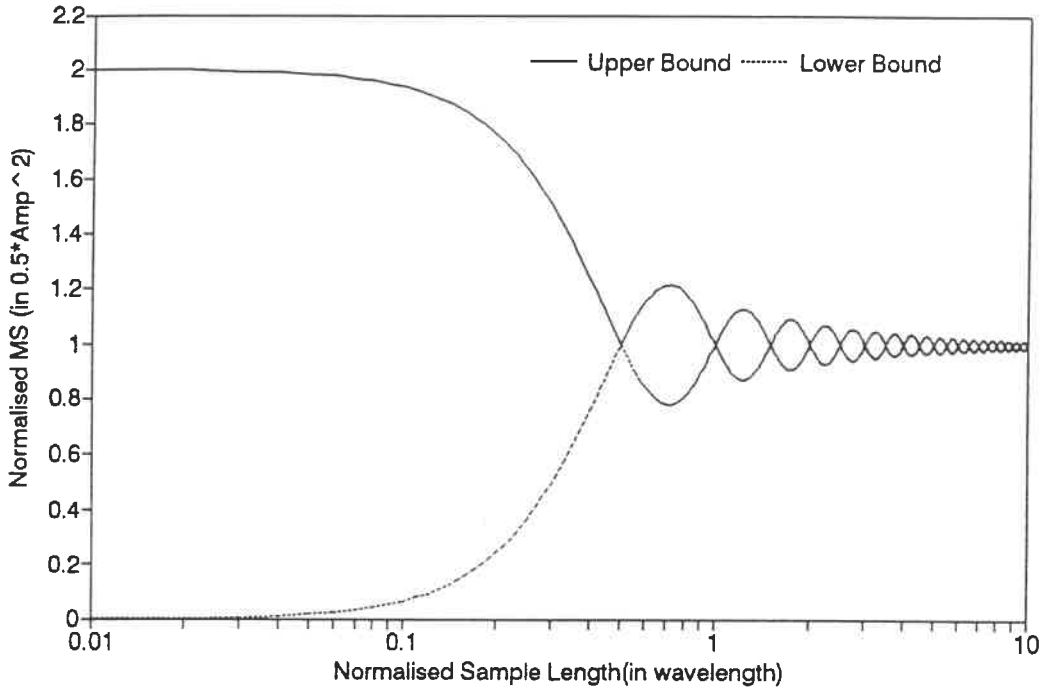


Fig. 1 Bound limits of mean square values sampled from a sinusoidal surface

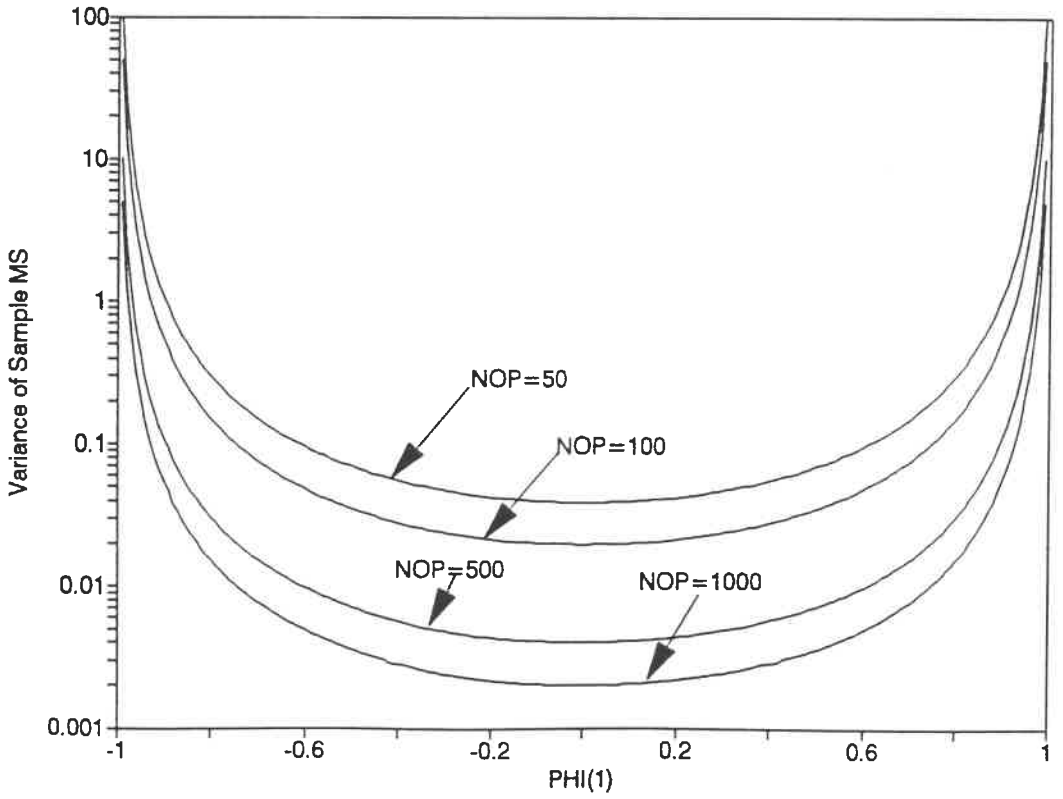


Fig. 2 Effects of correlation and sample size on the normalised variance of sample mean square (in terms of σ_0^4) from AR(1) surfaces

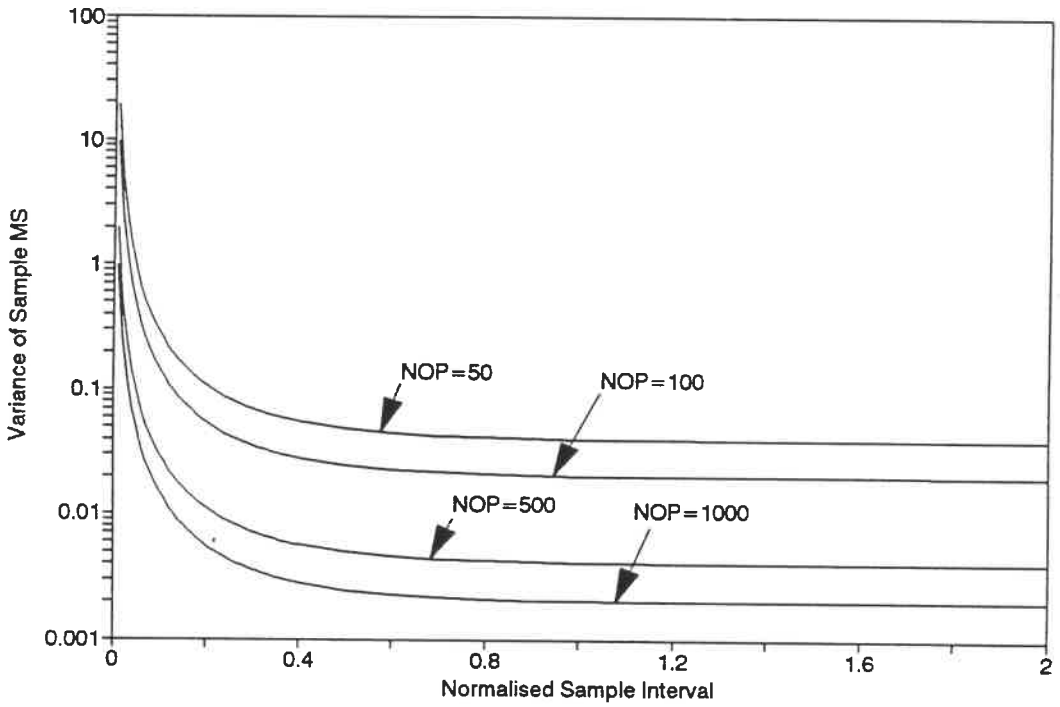


Fig. 3 Influence of normalised sample interval (Δ/β) on the normalised variance of mean square (in terms of σ_0^4) from AR(1) surfaces

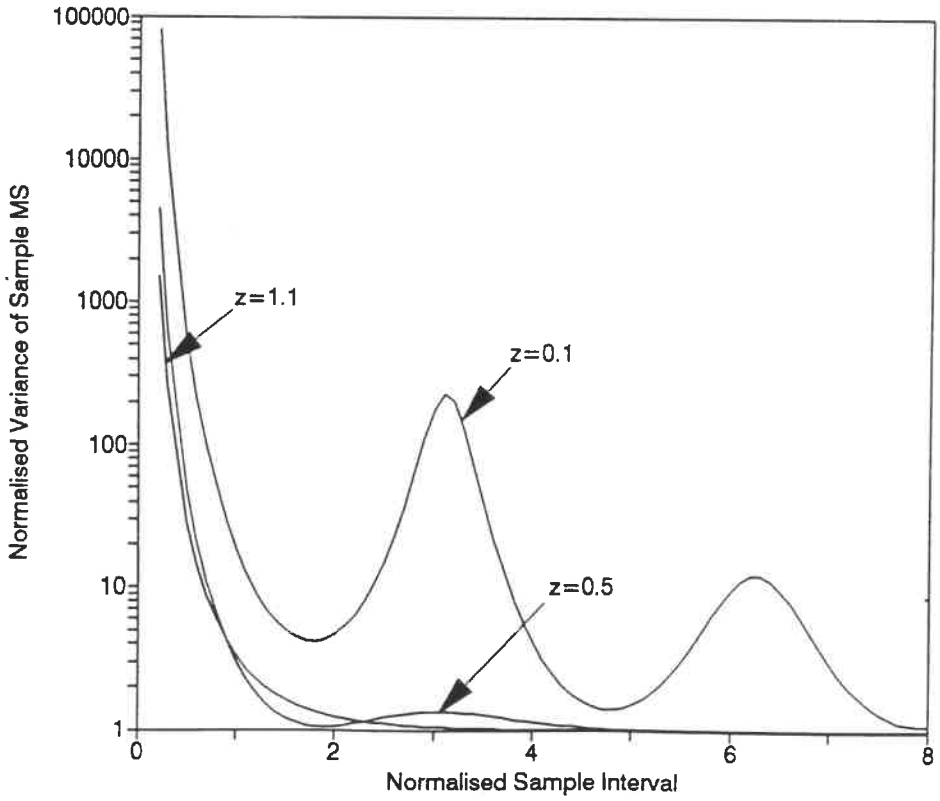


Fig. 4 Influence of normalised sample interval ($\omega_n\Delta$) on the normalised variance of mean square (in terms of $2(n-1)\sigma_0^4/n^2$) from AR(2) surfaces having a variety of process parameter zeta (ζ)

SIMULTANEOUS MEASUREMENT OF SPINDLE ERROR MOTIONS: A SIMPLE AND ECONOMICAL METHOD

Senthil Kumaran, Teruo Fujii, and Robert Hocken
Precision Engineering Laboratory
University of North Carolina at Charlotte
Charlotte, NC 28223

Introduction

The concept of measuring spindle error motions using electric pickups, inductive and capacitive, was described by Tlusty [1] as early as 1959. In 1967, Bryan, et al. [2], presented a clear picture of axis of rotation errors and measuring methods using capacitance sensors. In the years that followed, research has led to the development of experimental instruments for spindle error analysis at different research facilities worldwide: at Lawrence Livermore National Laboratory [3], at the Cranfield Unit for Precision Engineering, Bedford, U.K. [4], and at the Shanghai Machine Tool Research Institute, Shanghai, P.R. China [5]. Despite this experimental work, no commercial instrument conforming to the ANSI/ASME B89.3.4M-1985 and ANSI/ASME B5.54-1991 standards is currently commercially available to analyze an axis of rotation. If such an instrument is to be commercially produced, it should be easy to use, capable of measuring errors in both fixed and rotating sensitive directions, affordable, and if possible, applicable to other measurements. We have built a prototype of an instrument, the UNCC Spindle Error Analyzer (UNCC-SEA) that meets these criteria.

Principle

The principle employed here was first proposed by Tlusty [1] for rotating sensitive direction measurements. Two gages, positioned at right angles, sense against a master test ball mounted in the spindle with an offset introduced using a wobble plate. The eccentric setting of the master ball provides the base circle to which the error motions are vectorially added in the case of rotating sensitive direction measurements [1]. For fixed sensitive direction measurements, the error signal from the sensor positioned in the nonsensitive direction will be suppressed so that the polar plot reveals errors only in the sensitive direction. This is not currently implemented.

UNCC-SEA Hardware

Our analyzer uses a wobble-plate-adjustable master ball within a three-gage nest (Figure 1) to analyze errors in both fixed and rotating sensitive directions. The nest is made of INVAR to reduce thermal expansion due to local temperature variations. Strength calculations were performed to ensure stiffness of the wobble-plate. (The instrument can possibly be expanded to a five-gage nest sensing against two master balls for measuring tilt motion.) Gages used are all of 1 kHz bandwidth, which will be increased to

10 kHz as new gages become available. An IBM compatible PC with an 80386 microprocessor forms the base for all operations.

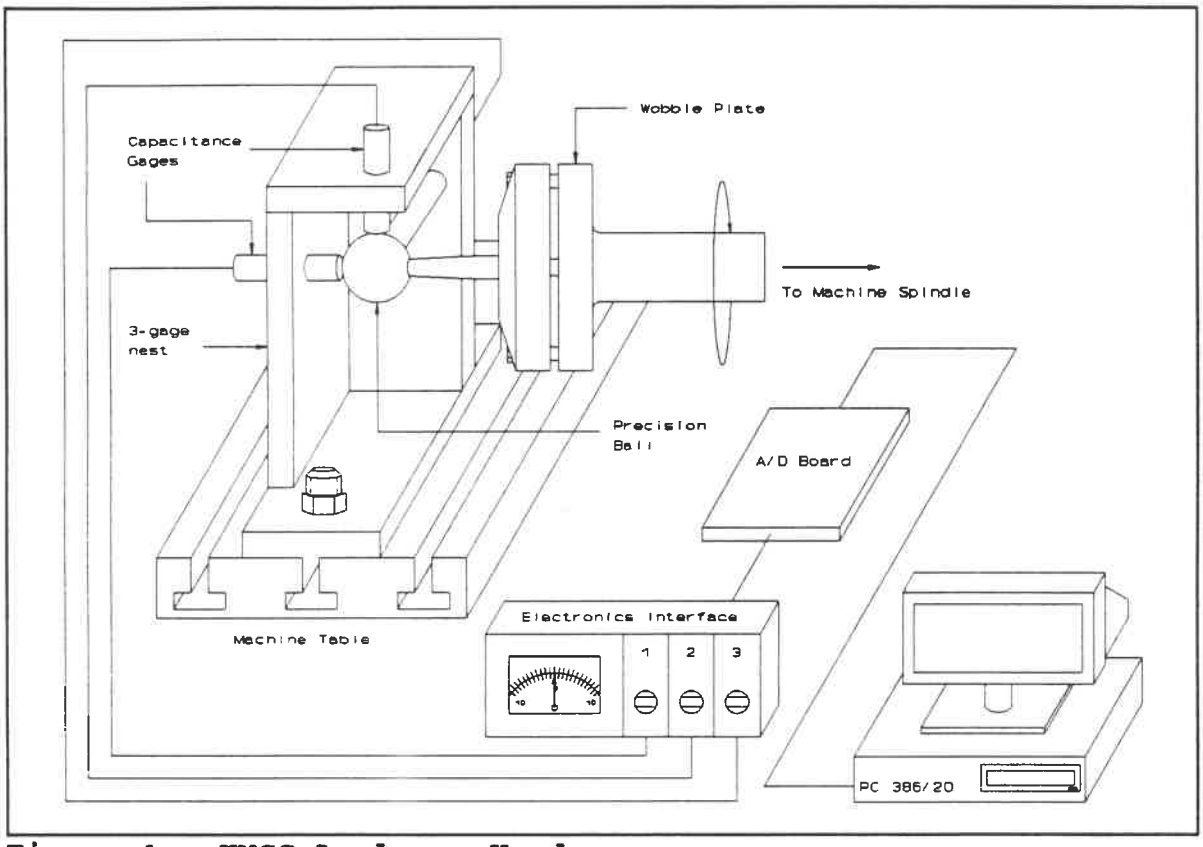


Figure 1. UNCC Analyzer Hardware

The master ball eccentricity is set at approximately 0.001 inch (about 30 μm; ten times the expected radial error motion value). The value of master ball eccentricity is of primary importance in obtaining a polar plot without any distortion [6]. The master test ball used here can be bought off the shelf and has a roundness error of approximately 5 μin. The gages used in radial motion measurement have a nominal standoff distance of 0.006 inch and the gage used for the axial motion measurement has a nominal standoff distance of 0.030 inch. Initial positioning of the gages is made using nonconducting 0.005 inch shims. The gages are then manually adjusted to be at the approximate center of the measuring range and secured in place using nylon screws. Any D.C. offset introduced is removed later during data analysis.

Data Acquisition

An eight channel data acquisition board with a throughput of 50 kHz is used to acquire data. The actual acquisition is performed using two DMA (Direct Memory Access) channels ensuring gap-free transfers of data to memory without any loss of samples. Also, the 25 kHz per channel sampling frequency comfortably satisfies the Nyquist criterion. The sampling rate is reduced to 8 kHz per channel when

the system is used in the real time display mode (see below) due to the time required to plot data to the screen.

UNCC-SEA Software

The analyzer operation is controlled by software developed at the Precision Engineering Laboratory at UNC-Charlotte. The software "UNCC-SEA" consists of four main modules: 1. Real-time display, 2. Data acquisition, 3. Data analysis and polar plots and 4. Data storage and retrieval. It is completely menu-driven and will include on-line help.

Real-time display: In this mode the computer functions as an oscilloscope. It graphically displays the spindle motion in real-time (Figure 2) with no noticeable delay in a polar plot form. As the spindle is stepped through its entire speed range the polar plot distinctly changes shape whenever resonant frequencies are encountered. The effects of these harmonics on spindle error motions have already been presented by Bryan and Vanherck [6]. These particular resonant frequencies can be noted and avoided during future machining.

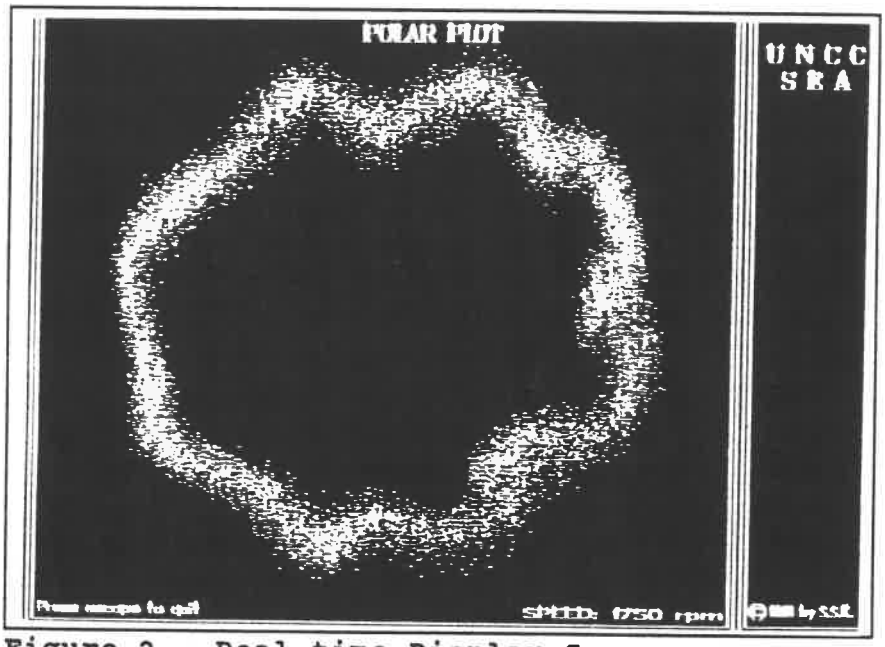


Figure 2. Real-time Display Screen

Data acquisition: The data acquisition is time based and more than 20 revolutions of data are collected. The sampling frequency is adjusted according to the signal frequency to maintain the number of samples per revolution a constant. Data for axial and radial error motions may be acquired and archived. At the time of this writing we are examining radial error only.

Data analysis and polar plots: A digital implementation of a Schmitt trigger with hysteresis (zero crossing detector) is used to obtain exactly 20 revolutions of data from the acquired raw data and thereby achieve data-acquisition synchronization [7]. This eliminates the need for an external trigger as described by Hansen [3]. D.C. offset, if present, is removed by subtracting the average from each data point and thus the wave forms are centered.

Then the Minimum Radial Separation (MRS) or the MiniMax center is computed. Using this MRS center, the asynchronous and average error motion polar plots are plotted and the values calculated, displayed and stored.

Data storage and retrieval: Raw data along with the calculated chart center and error motions can be stored to the hard disk or a floppy disk. Also any data file can be retrieved and analyzed.

Conclusion

For an otherwise perfect machine tool, from the polar plot and the error motion values, part roundness and surface finish potential can be predicted [8]. Bearing and structural error motions can also be diagnosed. Apart from achieving rotating sensitive direction measurements, the analyzer is also able to perform relative vibration tests and may be useful for temperature variation error (TVE), spindle thermal stability, multi-function cycle check, and repeatability tests on machining centers in accordance with ANSI/ASME B5.54-1991. It may also be usable for the machine compliance and hysteresis test if the load cell and fixturing are sufficiently stiff.

Acknowledgement

We would like to thank Lion Precision of Minneapolis for providing the capacitance gages and electronics.

References

1. Tlusty J., "System and Methods of Testing Machine Tools", Microtechnic, vol 13 (1959).

2. Bryan J., Clouser R., and Holland E., "Spindle Accuracy", American Machinist, December 4, 1967.

3. Hansen H., "A New Dynamic Spindle Analyzer", Proceedings of ASPE, Atlanta, 1988.

4. Chapman P., "A Capacitance Based Ultra-Precision Spindle Error Analyzer", Precision Engineering, 1985, p 129-137.

5. Su Yao-Sun, "A New Instrument For Axis of Rotation Metrology", Proceedings of the 7th SME International Precision Machining and Gaging Symposium, Cambridge, Massachusetts, 1986.

6. CIRP Unification Document, Axes of Rotation, CIRP Annals, 1978.

7. "Methods For Performance Evaluation Of Computer Numerically Controlled Machining Centers", ANSI/ASME B5.54, January 1991.

8. "Axes of Rotation, Methods for Specifying and Testing", ANSI/ASME B89.3.4M, 1985.

USING SCANNING PROBE MICROSCOPES FOR EXAMINING PRECISION MACHINED TOOLS AND MACHINED SURFACES

Mark R. Rodgers
Virgil B. Elings
Digital Instruments, Inc.
6780 Cortona Drive
Santa Barbara, CA 93117
Tel: 805-969-8116
Fax: 805-968-6627

Abstract

Scanning probe microscopy, SPM, refers to a class of instruments which provides very high resolution, three dimensional surface characteristics on a variety of samples. Two of the instruments in this class are the STM, scanning tunneling microscope, and the AFM, atomic force microscope. Although other scanning techniques exist, STMs and AFMs constitute the majority of scanning probe microscopes in use today. Both the STM and AFM operate in generally the same manner. While piezoelectric actuators perform high resolution raster scans in the X-Y plane parallel to the surface, a sensing and feedback system operates in the Z axis, perpendicular to the surface, measuring and controlling the variation in height across the sample. The main difference between the STM and the AFM is the sensing method used for the Z axis.

An STM relies on "tunneling current" between the probe and the sample for feedback on the topography of the sample. The STM probe, a sharp metal tip (in the best case, atomically sharp), is positioned a few atomic diameters above a conducting sample which is electrically biased with respect to the tip. At a distance under 1 nanometer, a tunneling current will flow from sample to tip. The exponential relationship between the tip separation and the tunneling current makes the tunneling current an excellent parameter for sensing the tip-to-sample separation.

An AFM uses a different method and a different feedback parameter to sense the topography of the sample. As with the STM, a piezoelectric actuator generates a precise scan in the X-Y plane. Most AFM techniques use the scanner to move the sample relative to a stationary tip. The probe is mounted on a flexible cantilever which deflects as the surface height of the sample changes. The deflection of the cantilever has been sensed with several techniques including tunneling current, optics, and capacitance. The AFM feedback technique is independent of the sample conductivity; therefore, unlike the STM, the AFM can image non-conductive samples. With new developments, the range of applications for SPMs continues to grow.

Our applications laboratory has received in recent months a variety of samples for evaluation which are of interest to the precision engineering community. The results of running the samples on SPMs will be presented. One type of sample is machined and formed optical surfaces, including glass and metal surfaces. We have also examined a variety of diamond turned surfaces. Precision machining tools have also been studied including diamond cutting surfaces and other fine cutting objects.

Using SPMs on these types of surfaces generate information that is not available from any other technology. The AFM, in particular, since it will image insulating surfaces directly, can provide resolution of fine surface details on a smaller scale than is practical with instruments that require conductive coating. Since data can be obtained in ambient conditions for many samples, the instruments are very simple to use.

Both STM and AFM data will be shown for several different applications. New developments that will increase the utility of these instruments will also be discussed. The following data will be shown at the poster session.

1. Micro-Machine Tool Wear-

A series of STM data (4) showing the abrasive effects of silicon. The diamond turning tool was used on gold both before and after being used on silicon. The prints show the degradation of the tool both qualitatively and quantitatively. This illustrates the use of the AFM/STM to analyze micro-machine tool wear.

2. Ytria-Stabilized Zirconia-

STM data showing Ytria-stabilized Zirconia grooved with a ruby ball. The large scan shows grooves while the higher resolution scan shows the details of deformation. Zirconia is an important substrate for a variety of applications including high T_c superconductors. The scan shows how it can be formed and machined.

3. Micro-Fabricated Mechanical Structure-

STM data of a Fresnel lens etched in HgCdTe. This lens is integrally fabricated to an element in a detector array. This image demonstrates the three dimensional imaging capability for microstructures.

4. Sub-micron Lines Etched in Gold-

Sub-micron lines etched in gold with an ion beam and imaged with an STM. This information is very simple and easy to acquire on sub-micron processes. This is done in ambient conditions, quickly and easily.

5. Polishing Defects in Hard Disk-

AFM image of polishing defects in a Canasite hard disk substrate. Demonstrates how the AFM can be used to evaluate precision processes on an insulating surface.

6. Polishing Defect in a Laser Mirror-

A polishing defect in a high power laser mirror using a large sample AFM. The image was obtained from the full size mirror which was 6 inches in diameter. This demonstrates a revolutionary capability for imaging large sample insulators.

7. Holes in Gold Made by STM Tip-

This is a series of progressively smaller holes in a straight line, made in gold, using a diamond STM tip. This indicates the use of an STM as a nano-fabrication tool.

8. Micro-holes Made by AFM Tip-

Very small holes made in a controlled pattern on an insulating surface using an AFM. The surface is BaBrF.

In addition, we will show data of general interest showing the current state of the art in:

- Large Scans
- Improved Tip Shape
- Electrochemical Operation
- Atomic Resolution
- New Capability of Large Sample AFM

Summary

1. The latest STMs and AFMs are scanning probe microscopes which can resolve surface detail down to the atomic level. This powerful high resolution capability has many present as well as future applications.
The size of scans ranges from areas as small as a few atoms to over 100 microns using a wide selection of interchangeable heads.

2. The capability of scanning probe microscopy has a large range of applications to Precision Engineering. By use of the AFM, non-conducting samples such as

ceramics, polymers, optical surfaces, and biological structures can be imaged directly. There is no need for coating, thinning, or other complex sample preparation.

3. In addition to SPMs' ability to image objects on the atomic and molecular scale, the SPM also emerges as a singular device for mapping and measuring three-dimensional surface profiles of objects in the range from 0.01 to 100 microns. The SPM can maintain <1 nm vertical resolution over this range.

4. Both the STM and the AFM are useful tools for monitoring and evaluating critical manufacturing processes. Manipulative processes are now possible and being used as in microelectronic mask repair.

Comparison of angle measuring interferometers and autocollimators

Paul Greier and Norman Bobroff
IBM Thomas J. Watson Research Ctr.
PO Box 218, Yorktown Hts, NY 10598

W. Tyler Estler and Yun Hsia Queen
NIST Gaithersburg, Md 20899

One method of assessing the accuracy of angle measuring autocollimators and interferometers is to compare simultaneous observations of a common target. The apparatus of Fig. 1 was used to compare Moller-Wedel Elcomat-2000 and Elcomat-HR autocollimators with Zygo Axiom and Hewlett-Packard 5527 laser interferometer systems. The instruments view separate faces of an octagon mounted on a rotary table. The table has a minimum angular increment of about 0.18″. The table steps are uncertain to about 0.02″ and absolute calibration is not presently possible. The HP and Zygo interferometer systems were configured to use the angle measurement feature of the Zygo 2001 combined angle/distance measuring optic. Because only one of these optics was available to us, separate data runs were taken for Zygo and HP interferometers. A conventional HP 10780 receiver was field modified to accept fiber optic input from the angle/distance optic. The resolution for the Elcomat-2000 is 0.05″, 0.005″ for the Elcomat-HR, 0.02″ for the Zygo axiom, and 0.08″ for the HP 5527. The resolution of the Elcomat-2000 and the interferometer system was extended by averaging readings. Sub count interpolation for the interferometers was achieved by averaging 200 measurements over an interval of about three seconds. Air turbulence adds a measurement noise of about +/-0.05″ to the averaged data. Nonlinearity of the interferometer systems was measured by observing the amplitude modulation at the receiver photodiode. No modulation was observed at a level of 1mV in 300mV, or about 1/2000 of a fringe. This is expected because of the complete symmetry of beam paths in the angle/displacement optic.

Measurement results are presented in Figs. 2-4. The vertical axis of Fig. 2 shows the angles measured by the Zygo Axiom and Elcomat-HR as the rotary table is moved through an angular range of about 22″. The measured angles are given relative to the nominal table position assuming a fixed step size of 0.18″. Because the actual table increments are variable about the nominal 0.18″ the data do not fall on a horizontal line. The scatter of the data points does not indicate a significant periodic contribution and we attribute the noise to air temperature fluctuations. Fig. 3 compares the HP 5527 system to the Elcomat-HR on the same scale as Fig. 2. Figs. 2 and 3 show the two interferometer systems to have comparable performance, with no evident systematic errors.

Data averaging at each table step is particularly favorable to the HP system which has a digital quantization limit of about 0.08″. The Elcomat data is slightly less noisy than the interferometer data and we attribute this to the shorter air path for the autocollimators. An interesting feature of these two figures is that the Elcomat-HR data agrees with the interferometer data at the start of the table motion before quickly falling to an offset about 0.07″ below the interferometers. This behavior of the Elcomat-HR was also observed in the comparison with the Elcomat-2000 and we have no explanation for it at this time.

Fig. 4 shows the Elcomat-2000 data for each of the two data runs. A systematic error having an amplitude of 0.15″ and a period of 3.6″ is evident in this data. The Elcomat-2000 uses a CCD detector to observe the return fiducial. The detection sensitivity of a CCD has a spatial variation which matches the pixel dimension of about 15-20um. This sensitivity variation modulates the fiducial signal, despite the fact that the fiducial image falls over several pixels. The period observed in the data is determined by both the pixel size and processing algorithm of this autocollimator which uses a multiple fiducial system. No such periodicity is observed in the Elcomat-HR data.

This simple experiment produced some unexpected results of significance to the precision engineering community. We hope to continue to characterize these systems at increased angular resolution with a new rotary encoder. Absolute calibration of the rotary table is also planned which will allow direct accuracy calibrations.

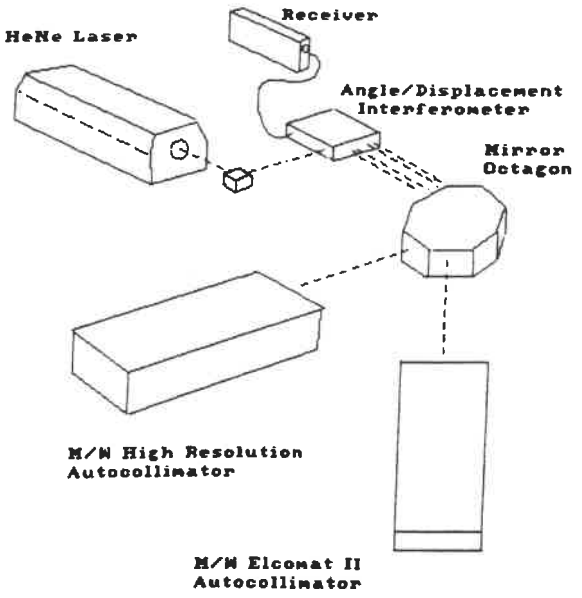


Figure 1. Angle Measurement System Configuration

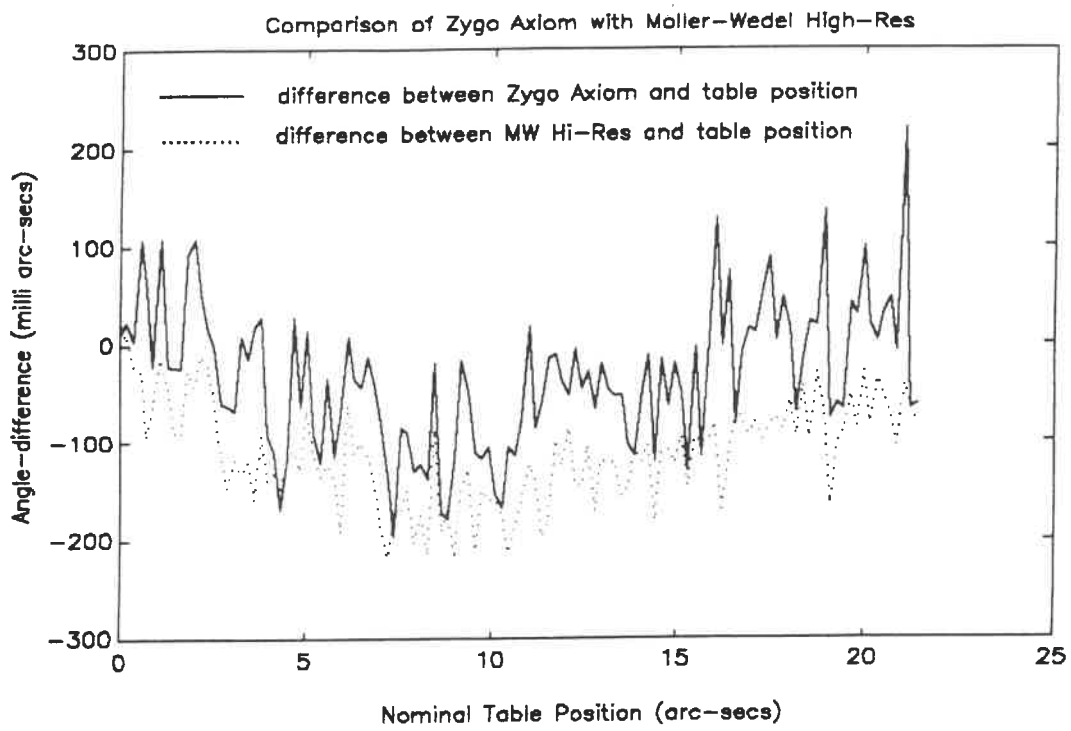


Figure 2. Zygo Axiom vs. Moller-Wedel High-Res Autocollimator

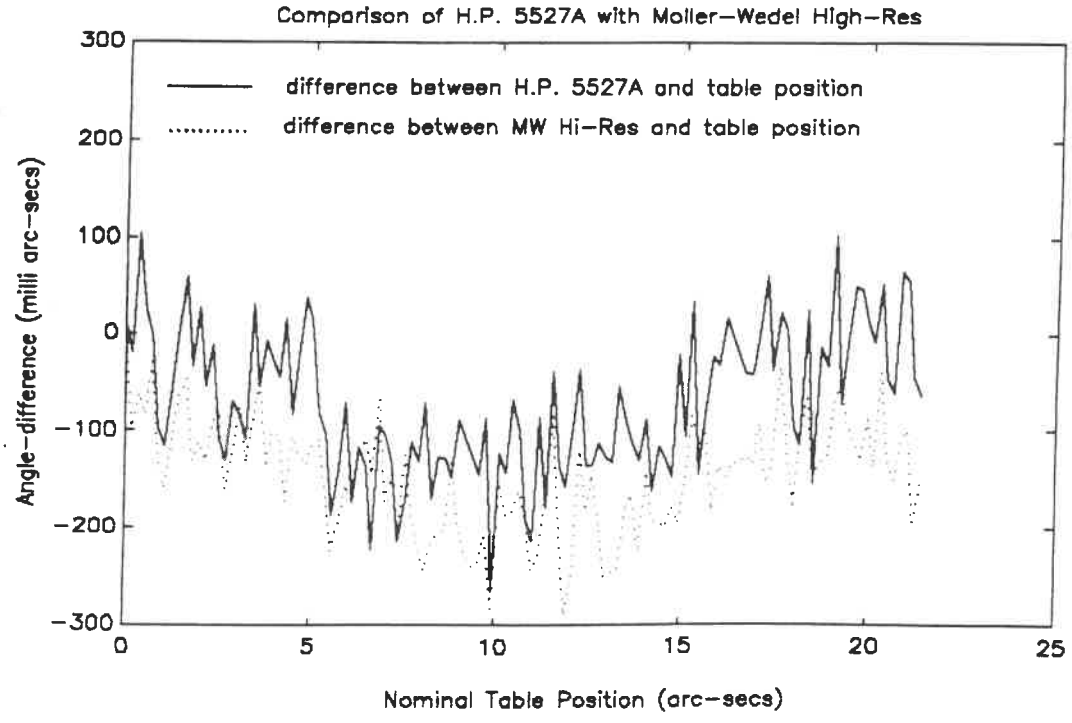


Figure 3. H.P 5527A vs. Moller-Wedel High-Res Autocollimator

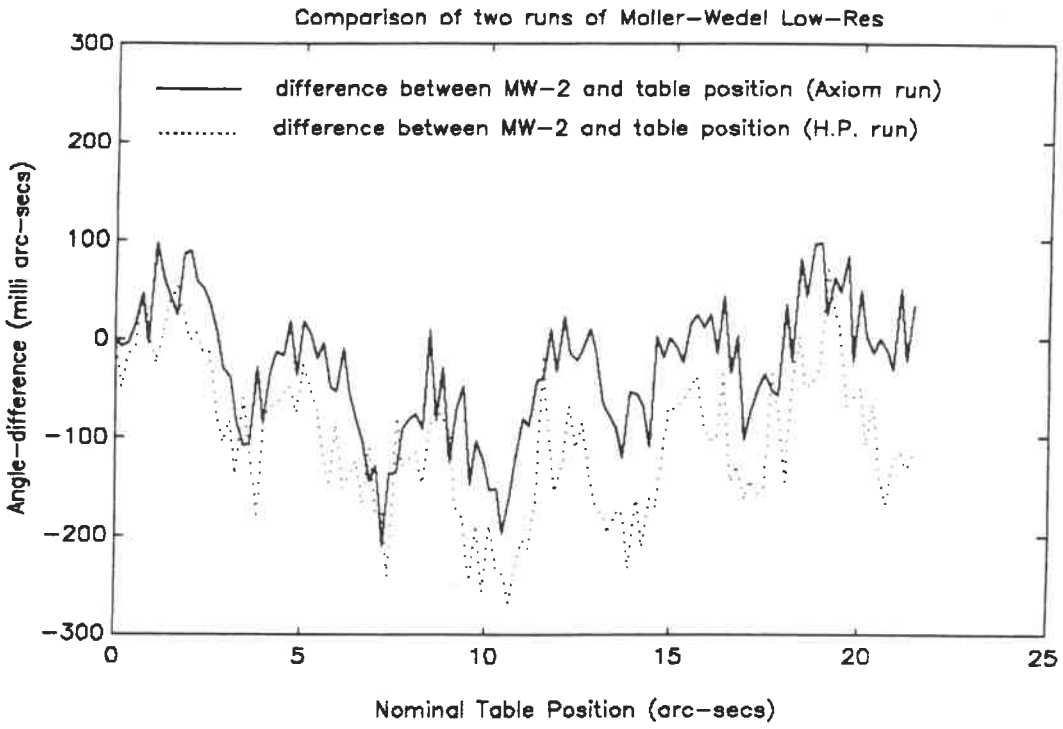


Figure 4. Moller-Wedel Elcomat II Autocollimator Data

THE CHARACTERIZATION OF MULTIPROCESS SURFACES (IS THIS A CASE FOR A PARADIGM SHIFT?)

Dr Paul J. Scott
Rank Taylor Hobson Ltd
Leicester, UK

1.0 Introduction

An ISO working group is currently investigating the characterization of multiprocess surfaces. The drive to characterize multiprocess surfaces is being lead by the automobile industry, a particular application being cylinder liners for engines. Their specification appears to require a surface that is cheap to manufacture that requires a minimum period to run in and has good bearing and lubrication properties (including a good distribution of oil reservoirs). This appears to be met by a plateau surface that has a number of evenly distributed deep grooves in it.

This is a very specific solution and is an attempt to optimise the surface texture for a particular function. It is also an example of a surface that has stratified functional properties. That is to say the higher parts of the surface (the plateau) have a different function to the lower parts of the surface (the deep valleys). Traditional analytical methods do not separate out these different functional components from the surface before characterization, resulting in a loss of information. In the next section several new characterization methods that address this problem are discussed together with their advantages and disadvantages.

2.0 Characterization Approaches

Five approaches to the characterization of measured profiles from a multiprocess surface are currently being considered, namely:

A/ Traditional Analysis - Profile filter, Characterization via traditional profile parameters (Reference [1]).

B/ R_k and friends - R_k Filter, Analysis on material ratio curve based on the minimum difference in height between points 40% apart (Reference [2]).

C/ Modified R_k - R_k Filter, Material ratio curve drawn on probability paper, analysis based on fitting two straight lines (Reference [3]).

D/ Motif method - Envelope filter, analysis based on R & W method (Reference [4]).

E/ Separation of plateau & valleys - Pattern recognition to separate plateau and valleys, analysis on separate constituents (Reference [5]).

2.1 Profile Filters

Traditional profile filters (such as the Gaussian filter), used in approach A, distort the profile. The deep valleys affect the filter and introduce artificial features into the profile around these valleys that are not part of the original profile. This leads to a distortion of the results.

The R_k filter (used in methods B & C) was an attempt to reduce this distortion. It was designed to be robust against deep valleys, it was partially successful though it still leads to some distortion (Reference [5]).

Methods D & E do not employ traditional profile filters to establish a reference line. They use a simple pattern recognition technique (Motif combination) to extract the relevant portions of the profile from which a reference line can be extracted. This approach appears to suffer less problems that traditional filtering, though the process is generally less understood.

2.2 Height Information

Traditional profile analysis (method A) uses the material ratio curve to characterize the height information of the measured profile. Various parameters such as R_a, R_{sk}, R_{ku}, R_y etc. can be calculated from the material ratio curve and used in its characterization. No attempt is made to distinguish between the plateau and the deep valleys in the calculation of these parameters.

The material ratio curve is also used in methods B & C to extract height information from the filtered profile. In an attempt to separate the plateau component from the deep valley component various algorithms are used to establish a point on the material ratio curve above which the curve is considered to characterize the plateau component and below which the curve is considered to characterize the deep valley component. These two components can be further divided in order to separate other information. Method B, for example, establishes a second point in the plateau component of the material ratio curve above which it is claimed characterizes the running in of the surface.

Like method A, the characterization analysis used in method D also does not distinguish between the plateau and the deep valleys.

An alternate approach, used in method E, is to separate the plateau and the deep valleys in the profile itself. Here simple pattern recognition is used to identify the deep valleys and to separate them from the plateau. Motif combination is particularly useful in this respect, since the individual motifs are defined, it is straight forward to choose between the plateau motifs and the deep valley motifs and characterize these separately. The characterization could use traditional profile parameters or R & W analysis.

2.3 Spacing Information

Functionally it is also important to characterize spacing information contained in the measured profile. The distribution of the oil reservoir along the profile is critical to prevent adverse effects

such as scuffing etc. Hence it is important to characterise, amongst other things, the distribution of deep valleys along the profile.

Both methods A & D calculate average spacing parameters from the measured profile. Again there is no attempt to distinguish between the plateau and the deep valleys. These average parameters can not discriminate between a profile in which the deep valleys are clustered together with large gaps between the clusters and a profile in which the deep valleys are spaced more evenly. There could be a large difference functionally between these two profiles.

Methods B & C currently do not contain methods to characterize spacing information. In practice these are often supplemented with traditional spacing parameters (Method A).

By separating the plateau from the deep valleys (method E) it is straight forward to characterize the individual components as well as to characterize the spacing between these components eg the maximum distance between deep valleys (important for scuffing).

2.4 Summary table

METHOD	FILTER	HEIGHT	SPACING	OTHER
A Traditional	Deep valleys distort Filter	No discrimination between deep valleys and plateaus	No discrimination between deep valleys and plateaus	Parameters well established
B R_k and Friends	Improvement over method A, still some distortion reported	Characterization of deep valleys, plateau and plateau peaks	No characterization of spacing information	Instabilities have been reported from industry
C Modified R_k	Improvement over method A, still some distortion reported	Characterization via material ratio curve	No characterization of spacing information	Algorithms have not been published. Little practical experience.
D Motif Method	Envelope filter, little or no distortion from deep valleys	No discrimination between deep valleys and plateaus	No discrimination between deep valleys and plateaus	Non-uniqueness of solution still needs to be resolved.
E Separation of plateau & valleys	Pattern recognition, no distortion from deep valleys	Deep valleys and plateau characterized separately	Deep valleys and plateau characterized separately	Details of method to be resolved. No practical experience.

3.0 Personal Comment

The future may bring more functionally optimized textured surfaces, that is surfaces that have been deliberately textured, where the texture has been optimized for a particular function. Many of these surfaces will also have stratified functional properties. Traditional surface metrological techniques will also fail to characterize these surfaces adequately since they can not discriminate between the various functional regions.

The approach that demonstrates the most capability from the list above is that based on separating out the various functional regions using pattern recognition. There are many pattern

recognition techniques that can be applied to measured profiles. The method suggested above is based on motif combination. This has proved very flexible and is computationaly very efficient (Reference [6]). Another simple and efficient pattern recognition method that has also given promising results is based on wavelet transforms. Flexibility is achieved by optimizing the analyzing wavelet to the particular problem in hand.

A more sophisticated approach is the use of functional simulation using the measured profile. Many simple functions such as rolling a circle or sliding a pad over a profile can be simulated using motif combination (Reference [6]), illustrating the flexibility of motif combination.

Both pattern recognition and functional simulation (or possibly a combination of both) represent possible future paradigms for surface texture analysis.

4.0 Conclusions

Plateau surfaces with deep groves in them represent an economically important family of surfaces. It is clear that traditional metrology fails to characterize these surfaces adequately. Approaches based on pattern recognition to separate regions of a surface that have different functional properties represents a possible future paradigm for the analysis of surface topography.

References

[1] International Standard ISO 4287/1 - (1984);
Surface roughness - Terminology - Part 1: Surface and its parameters.

[2] German Standard DIN 4776;
Parameters R_k, R_{pk}, R_{vk}, M_{r1}, M_{r2} to describe the material component in the roughness profile.

[3] Trautwein, R.;
Characteristic values for determining and evaluating the surface of cylinder bores, Stuttgart:Mahle GmbH (1981).

[4] French Industrial Standard E00.14.015.N -Août 1983;
Etats geometriques de surface calcul des criteres de profil.

[5] Schneider, U.; Steckroth, A.; Rau,N. & Hübner, G.;
An approach to the evaluation of surface profiles by separating them into functionally different parts. *4th International Conference on Metrology and Properties of Engineering Surfaces*, Washington D.C., United States of America.

[6] Scott, P.J.;
The mathematics of motif combination and its use for functional simulation. *5th International Conference on Metrology and Properties of Engineering Surfaces*, Leicester, England.

On-machine measurement of surface profile by fiber optic heterodyne interferometry

Katsuaki Sakaya and Toshio Ohura
Industrial Research Institute of Ishikawa
Ro-1, Tomizu-machi, Kanazawa-shi 920-02, Japan

INTRODUCTION

Recent progress and spread of the single-point diamond turning require the on-machine measurement of surface profile. Therefor, many non-contact surface measurement methods have been developed and reported. Especially optical methods with laser have been proposed the most owing to their high sensitivity and response. Of these optical stylus methods with heterodyne interferometry require no caribration and are capable of reliable measurement.[1,2] However, in present methods it is necessary to set the specimens on the instruments and measuring ranges are very small, because their laser beams from the sources are led to the interferometers and detectors by prisms and mirrors. In this report a small laser interferometer for non-contact optical measurement of surface profile is proposed, which is developed to be applied to the on-machine measurement. The following sections will describe the principle of this method, basic performance and the results obtained from the on-machine measurement of surface profile and surface roughness.

PRINCIPLE

The developed instrument is shown in Fig.1. The measurement principle of the instrument is based on the optical stylus method using single beam heterodyne laser interferometry. HP's laser system is used for the laser source and the detector. The beams from the laser source pass through a beam compressor to reduce their diameters to 0.6mm and a polarization maintaining optical fiber(PMOF). The transmitted beams are divided by a polarizing beam splitter. Each beams pass through 1/4 wave plates and one is focused by a microscope objective on the measured surface, while another is incident on the mirror. The reflected beams are combined in the polarizing beam splitter and pass through a multimode optical fiber to the detector. If the focused beam is scanned along the datum, the height variations of scanned line are measured. The interferometer shown in Fig.2, is a optical skid type interferometer for the on-machine surface roughness measurement. The principle of

this instrument is similar to the typical skid method. The two frequency components of laser with different spot sizes act as the stylus and the skid.

BASIC PERFORMANCE

The vertical sensitivity of the instrument is 10nm, and the lateral resolution, the measurable height and inclination by different kinds of objectives are shown in Table 1. The stability of the instrument shown in Fig.3, was measured in the condition that the specimen was fixed. Long term stability is influenced by the environmental temperature for thermal retardation of the PMOF. But the instrument has the same stability for short time as one without optical fibers. Vibration error reduction of the interferometer for surface roughness was measured by vibrating the specimen in 10Hz and 100nm by PZT actuator. In result it was confirmed that this error was reduced by 20dB. To determine the repeatability and the accuracy of the instrument, the surface profile of a diamon turned brass cylinder which was moved by the air slide was measured. Double-scan data by the instrument and the result by a stylus instrument are given in Fig.4.

MEASUREMENT

To examine the possibility of the on-machine measurement by the instrument, the cylindrical profile of a diamond turned brass cylinder which was mounted on a machine tool was measured. This result is shown in Fig.5. Also the surface roughness profile of the same specimen measured by the optical skid type interferometer is shown in Fig.6. The result is compared with the result obtained by a stylus method and the respective maximum surface heights were about 0.15um.

CONCLUSION

A technique for the on-machine measurement of surface profile using optical fibers has been described. Its vertical sensitivity of 10nm and very large measurable height compared with other non-contact optical methods are obtained. The mearsuring accuracy has been demostrated better than 20nm by repeated measurements and comparison with the stylus method. It has been demonstrated that this technique is applicable to the on-machine measurement of surface profiles.

REFERENCES

1. G. E. Sommargren, "Optical heterodyne profilometry", Appl. Opt. 20, 610 (1981)
2. Y. Tanimura and K. Toyoda, "Non-contact Measuring Instrument of Micro-shapes with Laser Interferometer", J. Jpn. Soc. Prec. Engng. 50, 1617 (1984)

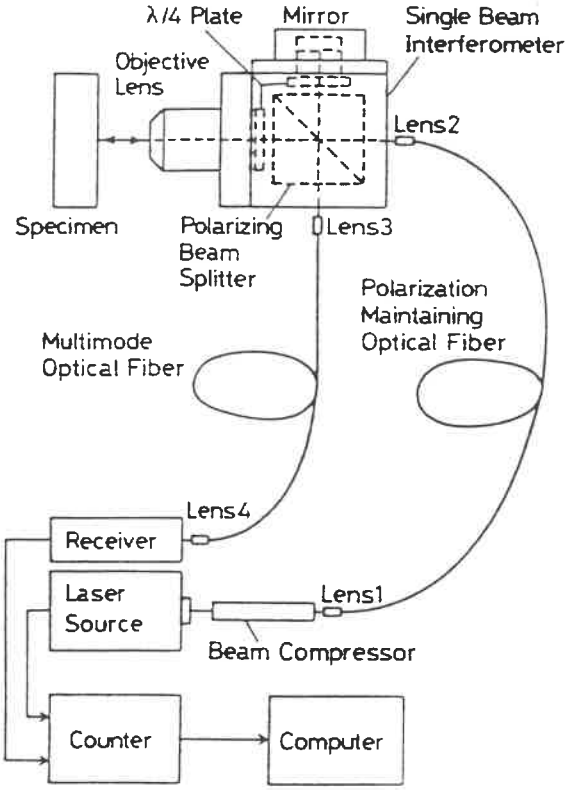


Fig.1 Schematic diagram of fiber optic heterodyne interferometer

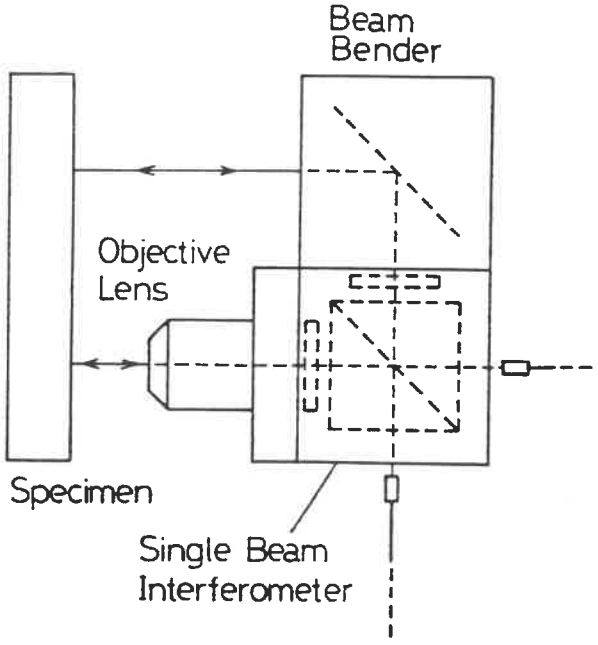


Fig.2 Schematic diagram of optical skid type interferometer

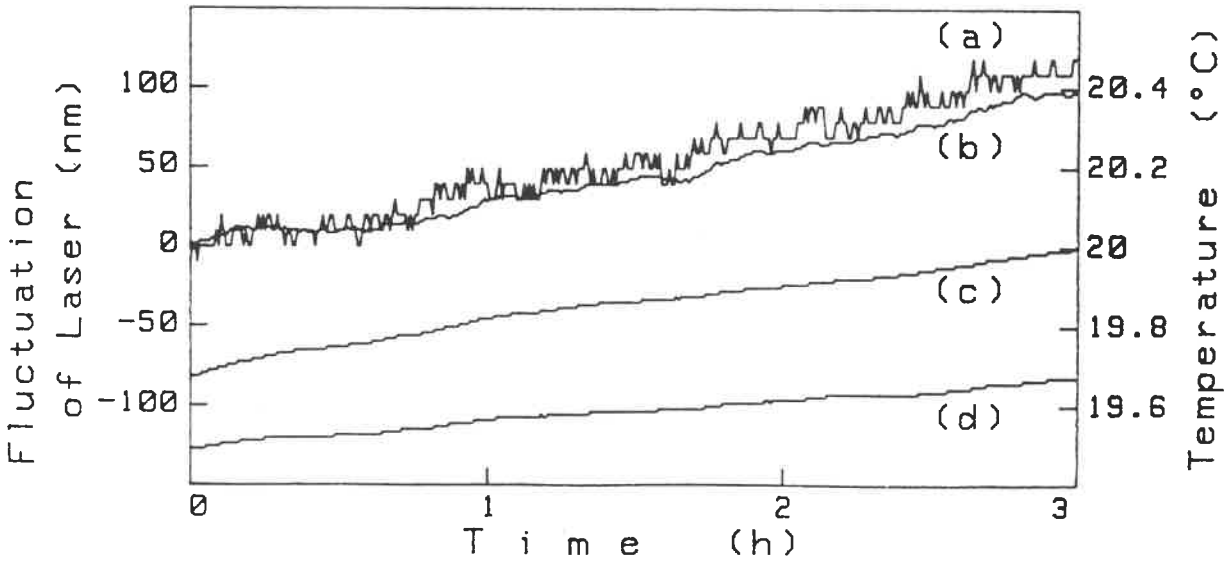


Fig.3 Long term stability of a displayed value of the interferometer (a)fluctuation of laser, (b)atmospheric temperature, (c)and(d)temperatures of surface plate

Table 1 Lateral resolution, measurable height and measurable inclination by different kinds of objective lenses

Objective lens (N.A.)	×10 (0.25)	×20 (0.40)	×40 (0.65)
Lateral resoluton (μm)	10.7	6.0	3.2
Measurable height (μm)	±550	±160	±40
Measurable inclination	±2.0°	±3.5°	±5.5°

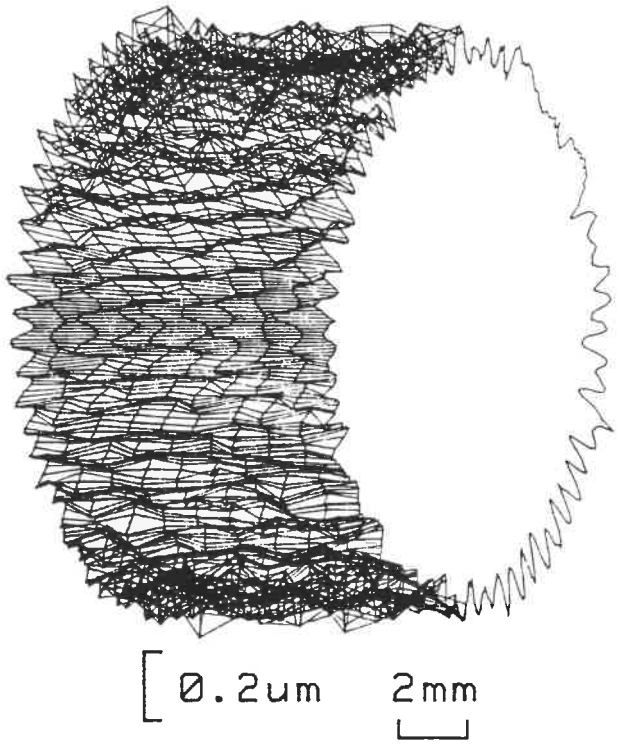
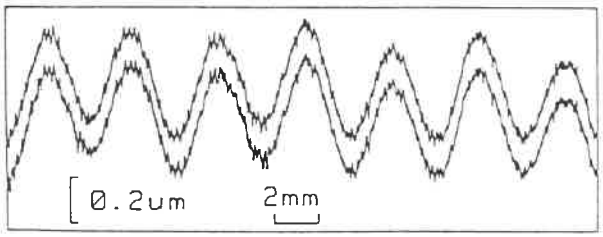
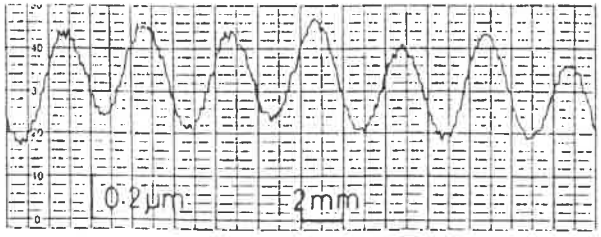


Fig.5 Surface profile of a diamond turned brass cylinder by on-machine measurement



(a) Double-scan data by the present instrument



(b) By a stylus instrument

Fig.4 Surface roughness profile of diamond turned cylinder

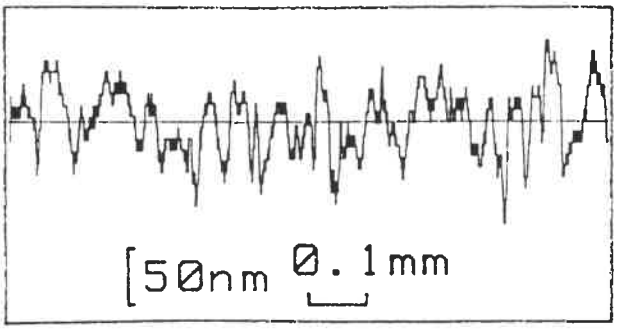


Fig.6 Surface roughness profile of a diamond turned cylinder by on-machine measurement

A HIERARCHICAL 2D PHASE UNWRAPPING STRATEGY IMMUNE TO SPIKE NOISE AND DISCONTINUITIES OF LARGER SCALE

T. R. Judge P. J. Bryanston-Cross

Engineering Department, Warwick University, Coventry, U. K.

There has been an ever increasing interest in the automation of Fringe Analysis over the last decade. The advent of Phase Stepping [1] and the development of FFT techniques [2] has moved the emphasis in Fringe Analysis away from fringe tracking and towards fringe counting or scanning techniques [3]. One difficulty with scanning approaches is their susceptibility to noise. Strategies have recently been proposed which exhibit noise immunity [4, 5], but which have failed to address large scale inconsistencies which cannot be detected at the pixel level, such as aliasing. Therefore, a hierarchical noise-immune technique has been developed which addresses both noise and large scale discontinuities.

KEYWORDS: interferometry, fringe analysis, graph theory, noise-immune phase unwrapping, tiling, minimum spanning tree

1 Introduction

A fringe field in general is a contour type map. However, the surface represented is sometimes discontinuous. This can occur for a variety of reasons; the object recorded could contain an edge, or the nyquist limit of the device capturing the interferogram could be exceeded which produces fringe aliasing. In addition to these discontinuity types, which can have broad influence in an interferogram, noise of various kinds, optical and electronic, will also be present.

Discontinuities can occur which have only a limited spatial extent and which can therefore be isolated to a small area of an interferogram. This concept is the basis for the development of an analysis strategy which separates the field into a number of segments (or tiles) in order to contain discontinuities.

The segmented field is then assembled under a priority scheme. Segments which seem to have solved unambiguously (when compared with their neighbours), may be employed in advance of less consistent ones. The priority scheme may be designed in such a way that it considers a variety of factors in drawing up this 'unwrap path'.

In Graph Theory [7, 8] , graphs are used to represent many types of problem. A classic example is the Traveling Salesman problem. That is of finding the shortest tour for a traveling salesman so that he visits each of a set of cities exactly once, and then returns to his initial city. The difficulty of this problem is illustrated by the immense number of possible tours: $\frac{(N-1)!}{2}$ for N cities. As this number grows faster than any

finite power of N, a brute force approach, that is a search of all candidates, is limited to small problems.

The TSP is a simple example of a difficult problem and is NP-complete (Non Polynomial Time Complete). The problem is analogous to several of more practical importance, including circuit layout and wire placement, resource allocation, and logistical problems. Research continues into algorithms which are able to produce suboptimal, but tolerable solutions in polynomial time.

However, the phase unwrapping problem is far more tractable. An efficient algorithm exists for the minimisation, upon which it relies.

The graph provides a structure in which a number of factors may be weighed. For a connected undirected weighted graph describing the unwrapping problem, a Minimum Spanning Tree may be used as a mechanism for comparing alternative phase unwrapping routes.

The graph approach is employed in a hierarchical strategy. At the lower level paths between pixels are compared to circumvent spike noise. At the higher level tiles of pixels are compared to circumvent the larger scale discontinuities. These are classified either as aliasing induced inconsistencies or natural inconsistencies.

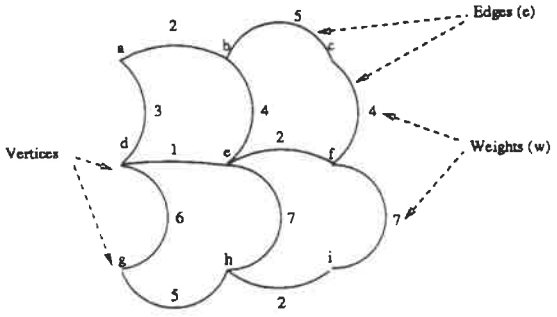


Figure 1: Example of a Weighted Undirected Connected Graph

2 Minimum Spanning Trees

A tree is a connected graph without any circuits. Figure 2 is a tree. Some properties of trees are outlined below.

There is one and only one path between every pair of vertices in a tree, T. If in a graph G there is one and only one path between every pair of vertices, G is a tree. A tree with n vertices has $n - 1$ edges.

A spanning tree in a graph G is a minimal subgraph connecting all the vertices of G. If graph G is a weighted graph (i.e., if there is a real number associated with each edge of G), then the weight of a spanning tree T of G is defined as the sum of the weights of all the branches in T. Among all of the spanning trees of G, those with the smallest weight are called Minimum Spanning Trees. Figure 2 shows the minimum spanning tree of the graph in Figure 1. [8]

3 Pixel Level Noise Immune Phase Unwrapping

Figure 3 illustrates pixel level phase unwrapping using the analogy of a mountaineer. The climber sets out from a plateau, in this case the foot of the mountain. He follows

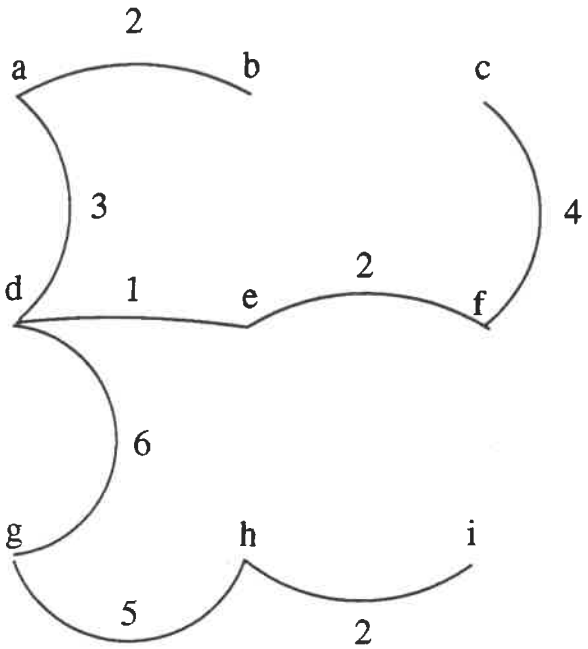


Figure 2: A Minimum Spanning Tree of the Weighted Graph

the easiest path he can find, where the ground is even and rises easily. The ruggedness of the terrain is similar to noise in the interferogram. When confronted by steeply rising ground, or an obstacle he cannot climb, he retraces his route and takes a less steep path.

4 Tile Level Phase Unwrapping

This level of phase unwrapping is aimed at dealing with larger scale inconsistencies which are not distinguishable at the pixel level.

A similar strategy applies to that adopted for coping with spike noise in that confidence is weighed by the minimum spanning tree. The aim is that areas of inconsistency should be contained within tiles. By comparing the unwrapped edges of neighbouring tiles, these contained inconsistencies may be detected and avoided in the unwrap path. The tiles are made to overlap slightly to facilitate this comparison.

5 Conclusion

The tiling/graph approach has produced an effective phase unwrapping strategy over a variety of discontinuity types. The technique is noise immune, solves around 'holes' and isolates localised discontinuities. The technique has been tested on a variety of cases, on both FFT and Phase Stepped maps [6].

The approach provides a structure for development. The graph weighting strategy may be easily augmented to accommodate new developments.

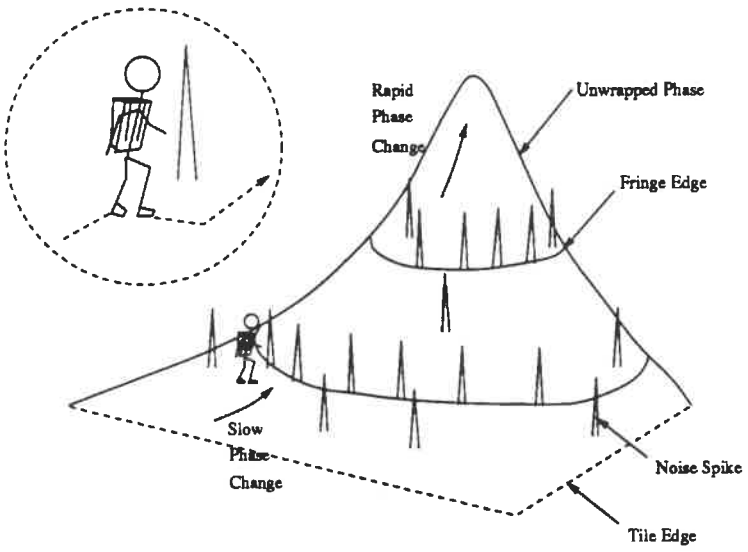


Figure 3: Mountaineer Analogy in Pixel Level Phase Unwrapping

References

[1] G. T. Reid, "Automatic Fringe Pattern Analysis: A Review," Optics and Lasers in Engineering, vol. 7, 37-68, (1986/7).

[2] M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," J. Opt. Soc. Am. 72, 156-160(1982).

[3] D. W. Robinson, and D. C. Williams, "Digital Phase Stepping Speckle Interferometry," Optics Communications, vol. 57(1), 26-30, (1986).

[4] D. C. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular-Automata Method For Phase Unwrapping," J. Opt. Soc. Am., vol. 4A, 267-280, (1987).

[5] J. M. Huntley, "Noise-immune Phase Unwrapping Algorithm," Applied Optics, vol. 28(15), 3268-3270, (1989).

[6] D. P. Towers, T. R. Judge, P. J. Bryanston-Cross, "Automatic Interferogram Analysis Techniques Applied to Quasi Heterodyne Holography and ESPI", invited paper to 'Optics and Lasers in Engineering', to be published.

[7] A. Gibbons, Algorithmic Graph Theory, 40-41, Cambridge University Press.

[8] N. Deo, Graph Theory with Applications to Engineering and Computer Science, 60-64, Prentice-Hall.

[9] R. C. Prim, "Shortest Connection Networks and Some Generalizations," Bell System Tech. Journal, vol. 36, 1389-1401, Nov. 1957.

[10] J. J. Gierloff, "Phase Unwrapping by Regions", 2-9, SPIE Vol. 818, Current Developments in Optical Engineering II (1987).

EVALUATION OF TOPOGRAPHY OF PRECISION ENGINEERING PART SURFACES

J. Mandak
Slovak Metrological Society
CZECHOSLOVAKIA

Growing demands on the quality of precision engineering part surfaces and automation of their production create new criteria for their accuracy, from the standpoint of their dimensions, shape and function. Prominent world producers of controlling and measuring instruments develope and manufacture, for this purpose, also measuring instruments for evaluation of topography - roughness - of the surface of machine parts. These devices usually work on electric, electromechanic, pneumatic, light or other principles.

The method based on the electromechanic principle has become most widely applied in the present day automized production of precision engineering part surfaces. This method employs an electric profile meter connected to a digital computer, e.g. 16 or 32bit microprocessor /measuring instruments of the trade mark FORM TALYSURF of the firm TAYLOR-HOBSON, SURFTEST III of the firm MITUTOYO, HOMMEL TESTER PM 2 and others/. These use various roughness characteristics for a quantitative description of precision engineering part surfaces /6, 7/ but none of them give adequate description of the topography - roughness - of the precision engineering part surfaces.

The method employing electromechanic principle uses scanning of the profile curve by a moving stylus. The profile curve is usually recorded on a memory medium in an enlarged scale, together with the set of roughness characteristics /6/ and usually as averaged value of i- successive measurements, together with the mean square deviation for the j-th numerical characteristics of roughness/6/.

The method employing electromechanical principle cannot be used at those places of a precision engineering part

surfaces where it is not possible to mount a controlling measuring instrument due to its mechanics and volume /e.g. openings with minimal diameter/. Prominent world producers of measuring instruments e.g. SURFTEST 211 of the firm MITUTOYO, try to minimize, at least partially, the above shortcomings.

In the given cases, where contact profile meters cannot be applied, light method can be used with advantage. This method may make use of virtues of optoelectronics, as e.g. multimode optical fibres, GaAs diodes, photodiode PIN, etc. These devices utilize the advantages of the stimulated emission of photons in an active laser medium /quantum electronic amplifier of electromagnetic radiation/ as a source of the monochromatic, coherent radiation, as to time and space, and the properties of the reflective diffusive surface of the machine part surfaces itself, compared with the applied light signal /determination of the size of the reflected light signal/.

Under conditions of the workplace there has been built an experimental measuring device of a module conception consisting of the power unit providing stabilized voltage to the exciting unit. This unit generates optical signal of a defined form /transformation of the electric signal into optical one/. It is created by GaAs diode, emitting pulse optical signal at the wavelenght of 900 nm. The evaluating unit consists of a detector with the photodiode PIN, designed to detecting reflectively modulated light signal and an active filter formed by an optimalized connection of Antony mutator. The light from the exciting unit is fed by the multimode optical fibre to the optical sensor that is placed in a special conductive casing, out of which the optical signal is emitted in a cone shape and an optical trace of a circle is being created over the reflective area of the precision engineering part surfaces. Its size depends on the present distance between the sensor and the illuminated area, on the type of the used optical fibre and on its numerical aperture. Portion of the reflected reflectively modulated light signal is conducted, in turn, by the multimode optical fibre to the evaluating unit. The transformed electric

signal is further transfered to the evaluating equipment. The advantages of the used multimode optical fibres are their low attenuation, large transfer information capacity, minimal size and weight, enhaced resistance against undesirable surrounding influences, etc.

The system consisting of one-chip microprocessor of the type 8035, intercepting register 3212 and program memory 2716 /EPROM/ was used for transfer of the electric signal to the evaluating equipment. The proper digitalization is carried out by the sampling amplifier MAC 198 and by the approximative A/D transducer WSH 570 that is connected to the system bus by a couple of three-state exciters 3216. USART of the type 8251 has been used for transforming 8-bit word into a series code.

The experimental measurements of roughness of precision engineering part surfaces performed with the device working on the electromechanic principle imply following conclusions:
- applicability of the light principle when compared to the electromechanic principle, can be seen only in the so called comparative measurement of roughness /it is necessary to have comparing etalons obtained by means of electromechanic principle/,
- the present day norms ČSN 01 4450 ST SEV 1156-78, or ISO /DIS 4287/1, DIN 4768/1 and others cannot be respected fully in the case of the light method, - the method based on the light principle has proved experimentally the influence of the radiation wavelength on the quantitative and qualitative results, - the quantitative results of measurements obtained by both procedures cannot be compared mutually to the full, - close attention has to be paid to the instantaneous state of the evaluated surface of the precision engineering part surfaces, above all to its cleanness.

Literature:

/1/ Manďák, J.: Possibilities of surface roughness evaluation

in automized manufacturing processes, IX. Days of Reliability - Assurance of Quality and Reliability in Automated Manufacturing Processes, Leipzig, FRG 1989.

/2/ Manďák, J. et al.: PV 4289-88, PV 4783-88, PV 4292-88, Inventions and Discoveries Office, Prague, 1988.

/3/ Deews, P. et al.: Quality assurance in the automated welding technique by the use of sensors, Technische Messen, n. 12, 1984.

/4/ Manďák, J. et al.: Quantification of surface micro-geometry with the use of He-Ne laser, International conference of Radioelectronics, Bratislava, CSFR, 1988.

/5/ Schmidt, B. et al.: Optical sensor technique, Feingerate-technik, n. 1, 1986.

/6/ Manďák, J. et al.: The influence of roughness of functional leading surface on the dynamics of friction tangential pair, Proceedings of the 3rd International Symposium Intertribo ´87, High Tatras, 1987.

/7/ Manďák, J.: Tribotechnical diagnostics of tangential pair with the use of quantification of changes of roughness characteristics, Technical Diagnostics 89, Prague, 198[illegible].

/8/ Manďák, J.: Evaluation of topography of machine part surfaces, Proceedings of the 4th International Symposium Intertribo ´90, High Tatras, 1990.

/9/ Manďák, J.: Evaluation of topography of machine part surfaces, Centro Argentino de Tribologia Comité técnico de la Cámara Argentina de Lubricantes, Buenos Aires, Argentina, 1991.

DEVELOPMENT OF A NEW VIBRATION REDUCTION CONTROL SCHEME FOR PRECISION SLIDE MOTION

Jeffrey A. Abler, Paul I. Ro
Precision Engineering Center
North Carolina State University

Introduction

Precision slide systems are often designed with a hydrostatic oil or air bearing to reduce friction. Because of the small amount of friction, the resulting system is highly underdamped in the direction of motion. This highly underdamped component makes the system extremely susceptible to vibration caused by the presence of external disturbances such as structural vibration, acoustic noise, and electrical noise. Given that a slide system is coupled to other components such as a motor and amplifier to drive the slide, there will be other dynamics in the system that are going to interact with the underdamped dynamics of the slide. This dynamic interaction can cause conventional filtering techniques to be ineffective and require a scheme which considers the dynamics of the entire system. A control scheme that dampens the oscillatory modes has been developed using a root locus design approach, and has been implemented on a slide system producing a significant reduction in vibration while also improving command following.

Design of the scheme was warranted because for a given set of motor/amplifier dynamics, and a given set of slide dynamics; efforts to improve each of these individually using conventional filtering techniques often oppose one another. The general problem encountered involves the competing factors of the amount of lead-lag added to the system to improve the dynamics of the motor and amplifier (improving system bandwidth and time response), versus the amount of lead-lag added to the system to dampen the lightly damped poles of the hydrostatic slide (reducing vibration). The improvement of one often has adverse effects on the other.

The design procedure developed separates the design of the motor/amplifier dynamics from that of the lightly damped slide poles. Once a suitable system bandwidth is established, then by placing a pair of zeros near the lightly damped poles of the slide, the poles are assigned an angle of departure which minimizes the sensitivity to disturbances while producing little effect on the system bandwidth. The procedure is outlined for the general case. Then, design and implementation on a Pnuemo ASG-2500 diamond turning machine is presented. Experimental results are compared to PID control and show a significant improvement in the level of vibration. Also, to demonstrate the effectiveness of the scheme in diamond turning, the roughness of machined surfaces are compared.

Design Procedure

Problem Definition

It is assumed that the discrete time transfer function of the N^{th} order system to be controlled can be expressed as

$$G(z) = \frac{K_G \prod_{i=1}^{N}(z - a_i)}{(z^2 - 2\sigma z + \sigma^2 + \omega^2)\prod_{j=1}^{N-2}(z - b_j)}$$

where $z^2 - 2\sigma z + \sigma^2 + \omega^2$ is a lightly damped pair of complex poles at $z = \sigma \pm i\omega$, which make the system susceptible to vibration. The remaining poles and zeros are more heavily damped, and one or more of these poles play a major role in determining the system's bandwidth.

The objective then, is to dampen the oscillatory poles while achieving the desired system bandwidth and time response. Classical lead-lag type

design can be used to achieve either of these goals, but often the system is such that it is difficult to improve one without diminishing the other. Other commonly suggested ways of dealing with undesirable dynamics include pole-zero cancellation; but it is widely understood that this is difficult to implement successfully because the location of the system poles cannot be known exactly, making such schemes sensitive to modeling errors and parameter variations.

Design Strategy

The strategy that will be employed here involves using a pair of complex zeros placed near the lightly damped poles not to cancel them, but rather to direct them to a more damped region. Because the zeros are not used to cancel the poles, the scheme will not be sensitive to variations or uncertainty in the location of the lightly damped poles (assuming the distance between the poles and the zeros is several times the uncertainty of the pole location). However, the distance between the zeros and the lightly damped poles will be small compared to the distance to the rest of the system poles. This will have the effect of making the rest of the system dynamics insensitive to the exact location of the complex zeros; because for poles far away, the effect of the complex zeros will cancel the effect of the lightly damped poles.

Since these lightly damped poles and zeros will have little effect on the rest of the system, design can first be carried out on the rest of the system ignoring the oscillatory poles. Once the controller has been designed to produce the desired system bandwidth and steady state characteristics, then the complex pair of zeros can be added to achieve the vibration reduction. This procedure is formalized below.

Formalization

The following steps are used in designing a discrete time "directional damping" controller utilizing root locus and frequency response techniques [1,2].

I. *Modify system $G(z)$ by replacing the lightly damped poles with two poles at the origin to form $G_{mod}(z)$.*

$$G_{mod}(z) = \frac{K_G \prod_{i=1}^{N}(z-a_i)}{z^2 \prod_{j=1}^{N-2}(z-b_j)}$$

These poles are at the origin in $G_{mod}(z)$ because a pair of poles will be placed at the origin when the complex zeros are added.

II. *Design forward loop controller $H_{mod}(z)$ for the modified system $G_{mod}(z)$ to achieve desired system bandwidth and steady state characteristics.* In this step, a lead-lag network can be used to increase the system bandwidth and assure that system and controller poles remain relatively far from the location of the lightly damped poles. Also, if desired, integral control can be incorporated to eliminate steady state following error.

III. *Add complex pair of zeros at the angle ϕ_z (see Figure 1) that will give the lightly damped poles the desired angle of departure to minimize sensitivity to disturbances.* The root locus angle criterion is used to determine the ϕ_z needed to produce the specified angle of departure. This criterion requires that for a system with transfer function $T(z)$ the root locus can exist only at the points z which satisfy

$$\angle T(z) = n180°, n = \pm 1, \pm 3, \pm 5, \ldots$$

Applying this near one of the lightly damped poles at a desired angle of departure ϕ_d, and using $n = 1$; the needed angular position of the complex zero is

$$\phi_z = 180° - \phi_d - \angle G_{mod}(\sigma + i\omega) H_{mod}(\sigma + i\omega)$$

In obtaining this result, the respective conjugates of the lightly damped pole and zero are ignored since they are at approximately the same angle with the pole, causing their effects to cancel one another. The complex zeros are then placed at

$$\alpha_{1,2} = (\sigma + r\cos\phi_z) \pm i(\omega + r\sin\phi_z)$$

forming the controller

$$H(z) = \frac{(z-\alpha_1)(z-\alpha_2)}{z^2} H_{mod}(z)$$

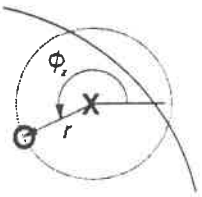


Figure 1: Complex Zero Placement Relative to Lightly Damped Pole

To minimize the sensitivity to disturbances, ϕ_d should be chosen to place the closed-loop pole location in direct line with the origin. This can be done by first choosing the angle of departure to be directly towards the center of the unit circle,

$$\phi_d = \frac{f_n T}{2} 180^\circ + 180^\circ$$

f_n = natural frequency of slide (Hz)

T = sampling rate (s)

and then modifying ϕ_d by the angle which the pole location deviates from the line to the origin.

The distance r between the poles and zeros should be several times larger than the uncertainty of the pole location to assure insensitivity to parameter variation and modeling error. However, if r is chosen too large the system bandwidth could be effected.

Implementation on Pnuemo ASG-2500 Diamond Turning Machine

Design

The scheme was implemented on a ballscrew driven, hydrostatic oil bearing slide way. The discrete root locus for the system using directional damping control with a $T = 1$ msec sampling rate is shown in Figure 2. The slide system has a natural frequency of 67 Hz which results in a complex pair of poles near the perimeter of the unit circle. These lightly damped poles make the system extremely susceptible to vibration, especially since there are many disturbances near this frequency.

A lead-lag network combined with a PID controller was used for H_{mod} to improve the bandwidth of the motor and amplifier while also producing zero steady state following error. An angle of departure of $\phi_d = 180^\circ$ was specified with a

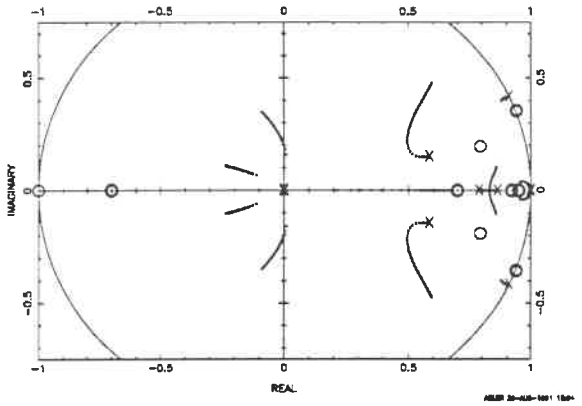


Figure 2: Root Locus of Slide System using Directional Damping Control. The bold X's and O's represent the controller poles and zeros.

pole-zero distance $r = .07$. Note that this ϕ_d requires a ϕ_z which introduces non-minimum phase zeros. This does not cause any stability problems since it would take much higher gains than will be used for the poles to track to the zeros, rather it is this positioning of the zeros which results in the motion of the poles to the most damped region.

The effect of the selected angle of departure on sensitivity to disturbances is shown in Figure 3. Smaller angles of departure result in less damped, higher frequency closed loop pole locations; while larger angles of departure result in less damped, lower frequency pole locations. This causes the peaks in the sensitivity if ϕ_d is varied by a large amount.

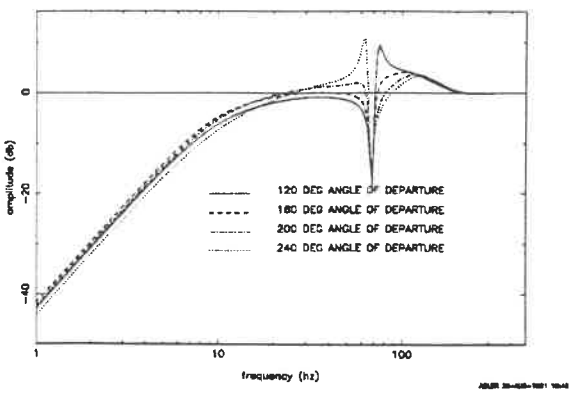


Figure 3: Sensitivity Frequency Response Using Directional Damping Control

Results

The performance of the control scheme was compared to PID control analytically using frequency response plots (of discrete model), and experimentally by measuring the vibration under each control at various feedrates. Plotted in Figure 4 is the sensitivity transfer function to disturbances after the plant such as acoustic noise and structural machine vibration. Directional damping control also greatly reduced the sensitivity to disturbances before the plant such as electrical noise.

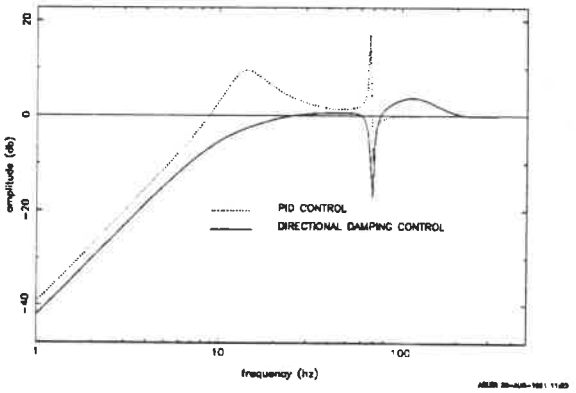


Figure 4: Sensitivity Frequency Response for Directional Damping and PID control

The measured rms vibration is plotted as a function of feedrate over the range of practical machining speeds in Figure 5. Directional damping control provided significant improvement in the level of vibration particularly when the slide is moving. The time response of the following error to a 5 mm/min move initiated at $t = .5$ sec is shown in Figure 6. The peak-to-valley vibration is improved from 250 nm to 100 nm. To demonstrate the effect of the vibration reduction on machined surfaces, sections of cones were diamond turned with the slide feeding at this rate. The peak-to-valley surface finish was reduced from 264 nm to 98 nm (49 nm to 17 nm rms) measured over a $780\mu m \times 570\ \mu m$ area.

Conclusion

A control scheme which separates the position loop design of the servo motor/amplifier dynamics

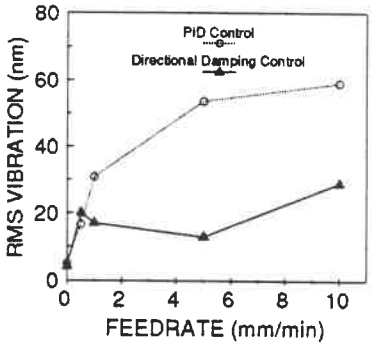


Figure 5: RMS Vibration for Directional Damping and PID Control

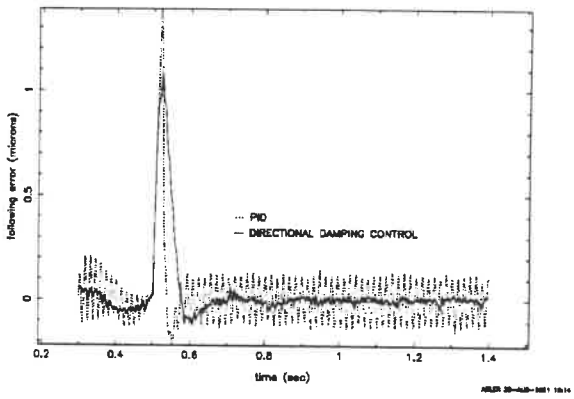


Figure 6: Following Error at 5 mm/min Feedrate

from the characteristically underdamped hydrostatic slide dynamics was developed using classical root locus and frequency response techniques. Implementation on the slide system of a diamond turning machine resulted in a significant improvement in the level of vibration over the range of practical machining feedrates. This reduction of vibration produced a corresponding improvement in the surface finish of machined parts.

References

[1] Evans, W. R., "Control System Synthesis by Root Locus Method", *AIEE Transactions*, vol. 69, part 1, 1950, pp. 66-69.

[2] Palm, W. J., *Modeling, Analysis and Control of Dynamic Systems*, Wiley & Sons, 1983.

THE PRELOADED BALL SCREW AS A NANOMETER ACTUATOR

James F. Cuttino
Thomas A. Dow
Precision Engineering Center
North Carolina State University
Raleigh, NC 27695-7918

INTRODUCTION

The field of precision engineering encompasses a wide range of high-accuracy manufacturing from semiconductor chips to meter sized optical assemblies. Fabrication of these components is possible with step-and-repeat cameras and diamond turning machines. The accuracy of such mechanisms, however, is highly dependent on linear slide mechanisms.[1] As emphasis is placed on more precise motion, the development of a linear drive system capable of moving a slide over large displacements with nanometer accuracy becomes an important topic of study.

This paper describes the characteristics of a ball-screw driven slideway known as the PLOT system (Precision Linear Optimization Testbed).[2] Due to the high-resolution required for many precision engineering applications, non-linearities in the slideway motion have become a topic of much interest. These non-linearities are discussed and a theoretical model is introduced to explain hysteretic losses.

LINEAR MOTION TESTBED

The testbed consists of a ball screw, a linear aerostatic slideway, and a non-influencing coupling connecting the drive system to the slideway as shown in Figure 1. For these small displacement experiments, the screw was activated using the ARC fixture, which consists of a flexure driven by either a piezoelectric actuator or a motorized micrometer. The PZT actuator was used for displacements of less than 100 nm and the motorized micrometer was used for those greater than 100 nm. The input angle was measured by a capacitance gage on the ARC fixture opposite the actuator. Slide displacement was measured using a pair of differential capacitance gages developed at Lawrence Livermore National Laboratories, which were calibrated with the ADE capacitance gage shown in Figure 1. Torque applied to the fixture was measured using a load cell in line with the PZT and the motorized micrometer.

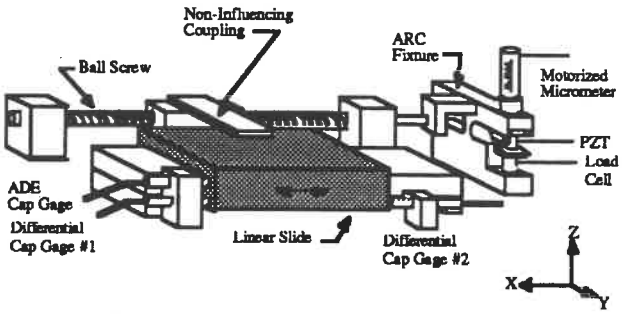


Figure 1: PLOT test fixture

Displacement Response to Differing Input Excitations Experiments were conducted to examine the ability of the PLOT to follow different types of inputs. These input signals included step input, ramp input, and sinusoidal input excitations. The step input experiment was conducted to examine the ability of the system to follow displacement commands of several nm. Figure 2 shows the slideway response to step inputs of approximately 3 nm over 2.5 second periods. The response of the system is quite impressive considering the open-loop nature of the system, with a noise level of approximately 2 nm for a mean step of 2.8 nm and a repeatability of less than 1 nm.

Figure 3 shows the slide's response to 120 nm ramp input. A nonlinear response is readily apparent and clearly indicates hysteretic effects have been introduced into the system. The deviation from a linear response is as high as 20 nm, corresponding to 17% error, and the system displays a higher slope towards the end of the cycle and returns to zero net displacement.

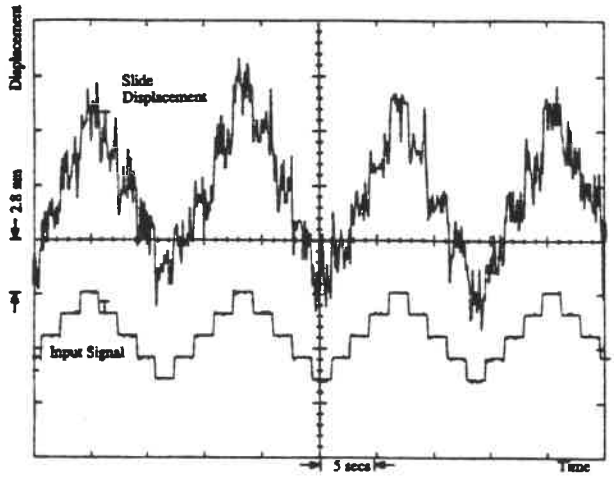


Figure 2. Slideway response to 3 nm input steps at 2.5 s intervals

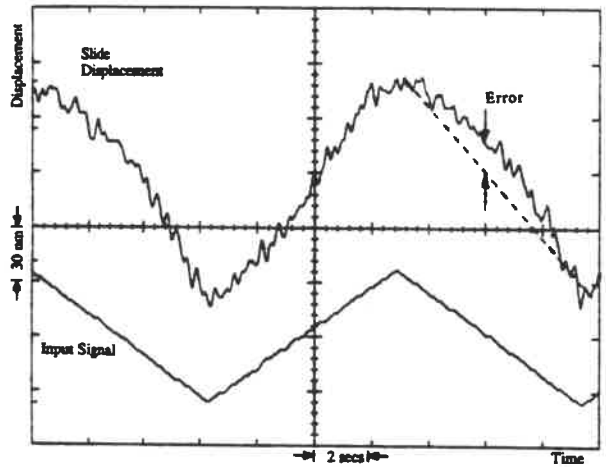


Figure 3: Displacement vs. Time Response for Ramp Input (120 nm ramp input over 12 s period)

Hysteresis Curves Figure 4 shows the measured displacement (solid curve) of the slideway as a function of the torque input. As torque is applied to the ball-screw, the slideway moves to the right as shown in Figure 4. When the maximum of 150 N-mm is reached, the torque is reduced and the ball-screw is driven in the opposite direction until the maximum torque is again reached. Note that the slope of the torque/displacement curve changes dramatically when the direction of motion is reversed.

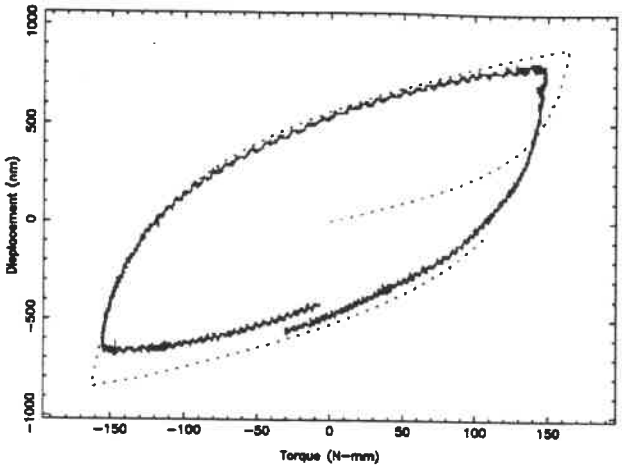


Figure 4. Ball-screw displacement vs. torque

Figure 5 illustrates the change in the measured pitch of the ball-screw (solid curve) as a function of input angle. For small angles (< 200 arcsec), the effective pitch of the system is significantly less than that of the screw itself. For motion more than 200 arcsec when the torque reaches its steady-state value, the system motion follows the pitch of the screw.

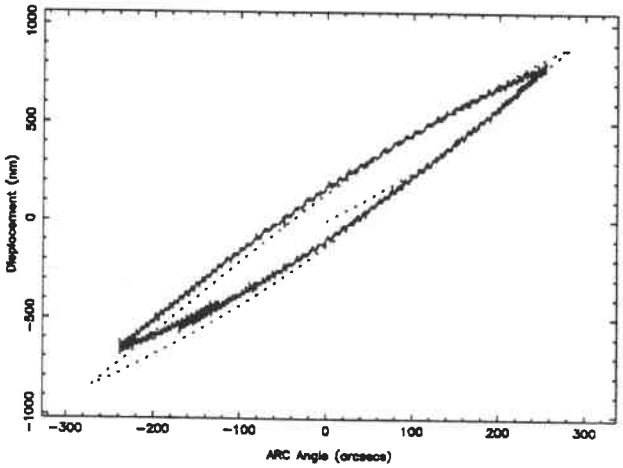


Figure 5. Ball-screw displacement vs. angular input

The system response to small angular inputs is a real problem for positioning slideways where sub-micrometer accuracy is required. In most instances, the ball-screw is assumed to be a simple gain in the control algorithm; that is, there is a fixed relationship between input angle and linear displacement. Figure 5 shows that

this is not true for displacements less than 200 nm and a model of this nonlinear behavior is the first step in controlling the mechanism over a range of input commands. This development is discussed in the next section.

MODEL OF BALL-SCREW FRICTION

The steady-state torque required to rotate the ball screw is assumed to originate in the forces necessary to roll/slide the balls between the stationary nut and the revolving screw. A combination of the elastic response of the screw and the deformation of the interface can result in the nonlinear response illustrated in Figures 4 and 5.

Ball Screw Geometry The interface pressure between the ball, nut, and screw is determined by the geometry of each component and the normal load. In this particular design, the diameter of the ball is larger than the clearance between the screw and the nut. This arrangement serves to preload the balls and to prevent backlash due to motion reversal. To reduce the friction inherent in such a design, every other ball has a reduced diameter and serves as a rolling spacer. Therefore, of the 80 balls in the working channel of the nut, only 40 are loaded.

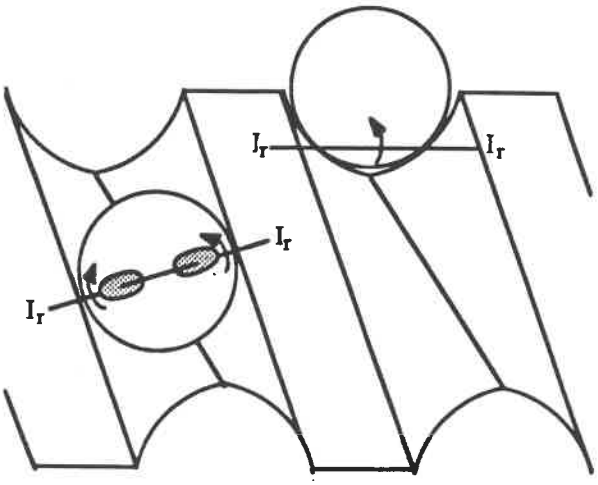


Figure 6. Contact between a ball and gothic arch groove in the ball-screw

The geometry of the ball screw contacts is shown in Figure 6. The screw and nut are formed with two large radii forming a "gothic arch". This larger radius reduces the peak pressure at each of the elastic contacts by increasing the area of contact.

Studies conducted by Ishikawa and Suda[3] refer to the frictional effects found for a ball rolling in a V-shaped groove. By modifying this approach and applying it to the gothic arch profile, the frictional force required to roll/slide the balls in the nut was calculated.

Referring to Figure 6, I_r is the instantaneous axes of rotation for the ball against the screw. The ball contact region over the gothic arch is shown. Since the ball rotates about an instantaneous center point, any other point in the contact patch must undergo rotation with respect to that center. The angular rotation of the contact is the same as that of the ball.

Once the pressure distribution and contact patch shape are known, the total torque of the ball may be calculated based on a summation of the contribution of sliding friction from each point within the contact area.

Contact Patch Geometry The component geometry and preload were obtained from the manufacturer or, where proprietary (such as arch radius), were measurements on the ball screw. The elastic contact areas between the ball and the nut and the ball and the screw were analyzed to find the pressure distributions and contact dimensions. These two contacts are different due to two factors. First, the contact with the nut is at a radius larger than that of the screw. Second, the curvature of the nut is negative, whereas the curvature of the screw is positive. This will mean that the contact between the ball and the nut will be larger than the contact between the ball and the screw.

Torque Calculation Based on the pressure distribution and contact dimensions, a program was written to discretize the contact into small areas, and, based on a given friction coefficient, to calculate the torque necessary to twist the contact about the instantaneous center of rotation shown in Figure 7.

The pressure distribution is elliptical over the contact and is maximum at the instantaneous center of rotation. The contribution of torque is

zero at the center, increases away from center until a maximum torque is reached, and then decreases to zero again around the outer edge of the contact. The calculated torque for an estimated 40 balls in contact yields a total torque of 170 N-mm. This value is very near the steady state torque measured for the PLOT apparatus.

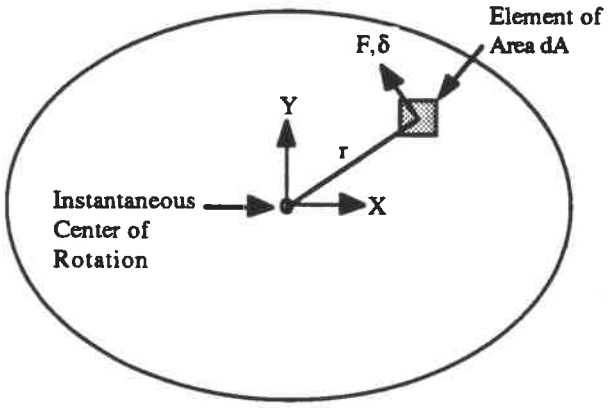


Figure 7. Contact Patch Discretization and Torque Calculation.

Non-Linear Motion The friction model described above can also be used to determine the origin of the non-linear relationship between input angle and output displacement. The torque required to roll the balls along the groove between the nut and the screw will produce elastic deflections at the interface surfaces as well as in the screw itself. The elastic windup of the screw can be calculated from the screw cross-section, length, and magnitude of the torque. The interface between the ball and nut/screw is more difficult to model. If this interface is assumed to be made up of small elastic elements that stick up from the surface, an expression for their response can be developed. Each element is allowed to deflect elastically until the load exceeds the normal force times the friction coefficient, beyond which the element will slide and the deformation is kept constant. With this model, the outer elements will slide sooner than those near the center of rotation as a consequence of the longer moment arm and the small normal load at the outside edge of the contact. This will result in gradual breakaway and a non-linear relationship between torque and rotation of the screw. This relationship coupled with the windup of the shaft can be used to predict the ball-screw performance.

Theoretical Results Two parameters are needed to model the performance of a ball-screw drive: the friction coefficient and the stiffness of the ball interface. The measured response was used to select these two values and resulted in a reasonable friction coefficient of 0.11 and an interface stiffness of 168 N/μm.

The predicted performance of the ball-screw drive is shown as the dotted lines in Figures 4 and 5. The predicted curves start at the origin of torque and displacement and the experimental curves have been centered with respect to the predicted values. When the drive is first loaded, all of the interface elements are theoretically unloaded. As a result of loading however, some residual elastic deflection is stored in the interface and the response to subsequent motion is quite different as illustrated in Figures 4 and 5.

CONCLUSIONS

The preloaded ball-screw can be used for nanometer resolution actuation; however, the response is not linear. The nonlinearity has been shown to be a consequence of elastic windup in the screw and deformation at the ball contacts. A model has been derived which predicts the friction and displacement as a function of input angle. It will serve as the basis to develop a control strategy to operate over a range of displacement commands.

REFERENCES

1. Futami, Shigeru, Akihiro Furutani, and Shoichiro Yoshida, "Nanometer Positioning and Its Micro-dynamics", *Nanotechnology*, Volume 1, No. 1, July 1990, pp. 31-37.
2. Knight, B. F., "Development of a Testbed for Precision Linear Motion", MS Thesis, Department of Mechanical and Aerospace Engineering, NCSU, May 1990.
3. Ishikawa, Yoshio, and Minoru Suda, "Differential Slipping Friction of a Ball Rolling in a V-Shaped Groove", *The Journal of JSLE*, International Edition, November 7,1986, pp. 29-34.

Design of Three Groove Kinematic Couplings

Alexander H. Slocum
Department of Mechanical Engineering
Massachusetts Institute of Technology

Introduction

Kinematic couplings have long been known to provide an economical and dependable method for attaining high repeatability in fixtures [1]. Properly designed kinematic couplings are deterministic: They only make contact at a number of points equal to the number of degrees-of-freedom that are to be restrained. Being deterministic makes performance predictable and also helps to reduce design and manufacturing costs [2]. On the other hand, contact stresses in kinematic couplings are often very high and no elastohydrodynamic lubrication layer exists between the elements that are in point contact; thus for high cycle applications, it is advantageous to have the contact surfaces made from corrosion resistant materials (e.g. ceramics). When non-stainless steel components are used, one must be wary of fretting at the contact interfaces so steel couplings should only be used for low cycle applications.

Coupling Configuration and Stability

Symmetry aids in reducing manufacturing costs, and for practical fixturing applications, in general, the use of grooves for all contact regions minimizes the overall stress state in the coupling. Thus it is assumed here that the kinematic coupling to be designed is a three groove type.

Two forms of three groove couplings are illustrated in Figure 1. Planar couplings are often found in metrology applications. They can also be used in the manufacture of precision parts. For example, a planer coupling can be used to hold a grinding fixture on a profile grinder. A matching three groove plate on a CMM allows the grinding fixture to be transferred to the CMM with the part. The part can be measured and then placed back onto the grinder so the errors can be corrected. To minimize Abbe errors in some applications, vertically oriented couplings can be designed where the preload is obtained with a clamping mechanism or by gravity acting on a mass held by a cantilevered arm. An example would be a three groove kinematic coupling used to hold photolithographic masks in a wafer stepper whose projection axis must be horizontal because of its size.

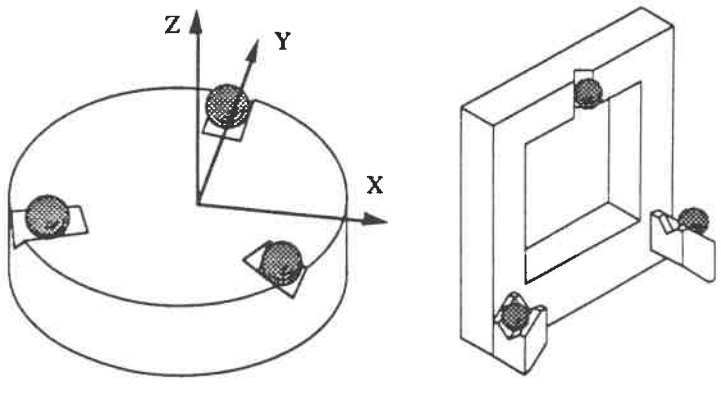


Figure 1 Examples of three groove kinematic couplings for horizontal and vertical fixturing applications. The coupling components to which the balls are permanently affixed are not shown for clarity.

With respect to practical implementation of the theoretical requirement for stability, for precision three groove kinematic couplings, stability and good overall stiffness will be obtained if the normals to the plane of the contact force vectors bisect the angles of the triangle formed by the hemispheres (e.g., balls) that lie in the grooves[1]. Furthermore, for balanced stiffness in all directions, the contact force vectors should intersect the plane of coupling action at an angle of 45 degrees. The angle bisector concept is illustrated in Figure 2. Note that the angle bisectors intersect at a point which is also the center of the circle which can be inscribed in the coupling triangle. This point is referred to as the *coupling centroid* and it is only coincident with the coupling triangle's centroid when the coupling triangle is an equilateral triangle.

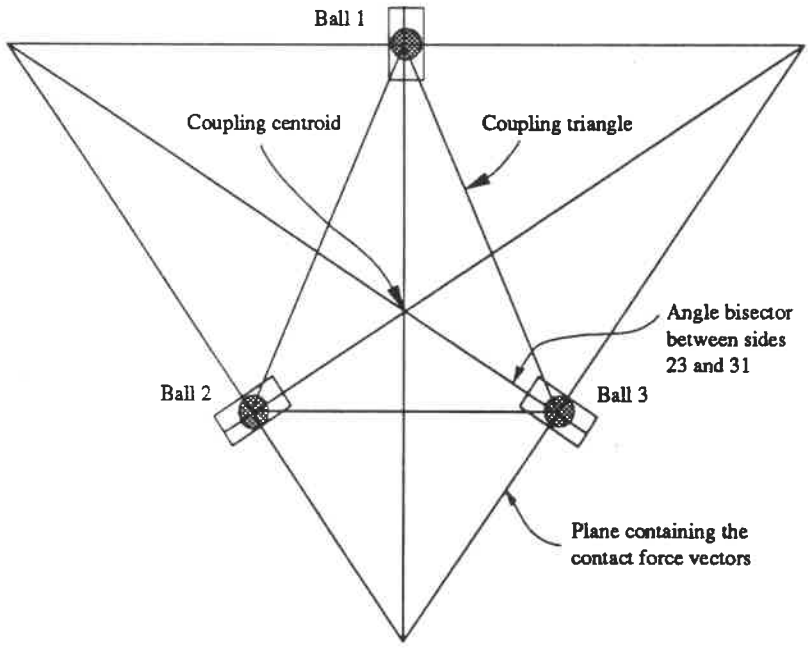


Figure 2 For good stability in a three groove kinematic coupling, the normals to the planes containing the contact force vectors should bisect the angles between the balls.

Note that any three groove kinematic coupling's stability can be quickly assessed by examining the intersections of the planes that contain the contact force vectors. For stability, the planes must form a triangle as illustrated in Figure 4.

---

[1] From conversations and observations with Dr. William Plummer, Director of Optical Engineering, Polaroid Corp, 38 Henry Street, Cambridge, MA 02139.

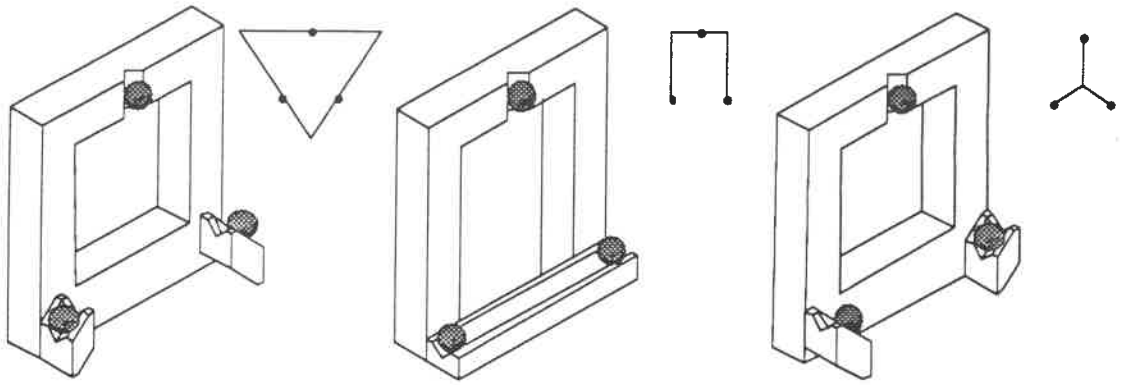


Figure 3 Different configurations for a kinematic coupling which illustrate how the intersections of the planes containing the contact force vectors can be used to make an assessment of the coupling's stability.

Kinematics of the Coupling's Error Motions

The contact between the ball and the groove actually results in an elastic indentation of the region. Combined with a finite coefficient of friction, it is reasonable to assume that there is no relative motion between the ball and the groove at the contact interface. If one makes this assumption and then calculates the new position of the balls' centers using the contact displacements and contact forces' direction cosines, then one finds that there is not a unique homogeneous transformation matrix that relates the old and new ball positions. These factors make the calculation of a kinematic coupling's error motions a non-deterministic problem.

If the distances between the balls, determined using their new coordinates, do not change greatly, then reasonable estimates can be made of the coupling's error motions. Using the design theory presented in this paper, a spreadsheet can be used to show that the change in distance between the balls is typically five to ten times less than the deflection at the contact points. Furthermore, the ratio of the change in the distance between the balls to the distance between the balls is typically an order of magnitude less than the ratio of the deflection of the ball to the ball diameter (see the calculations in Appendix A). Thus closed form analytical estimates of the coupling's error motions can be made.

Results and Conclusions

The analysis methods described above were implemented on a spreadsheet. In order to test the spreadsheet, forces were applied along axes of symmetry and it was checked to ensure that the expected displacements were obtained. For example, a z direction force should yield equal forces at all the contact points and only a z displacement should occur. For all test cases, the results were as expected.

References

[1] C. Evans, Precision Engineering: an Evolutionary View, Cranfield University Press, Cranfield England, 1989, pp 21-29.

[2] A. Slocum, "Kinematic Couplings for Precision Fixturing - Part I - Formulation of Design Parameters", Precision Engineering, Vol. 10, No. 2, April 1988, pp 85-91.

[3] A. Slocum and A. Donmez, "Kinematic Couplings for Precision Fixturing - Part II - Experimental Determination of Repeatability and Stiffness", Precision Engineering, Vol. 10, No. 3, July 1988, pp. 115-121.

A Fundamental Structure for Intelligent Manufacturing

Yotaro HATAMURA, Takaaki NAGAO and Mamoru MITSUISHI
The University of Tokyo
Hongo Bunkyo-ku, Tokyo 113, Japan

In order to realize a high precision, flexible manufacturing system with easy operation and 24 hour unmanned capability, a completely new concept for "Intelligent Manufacturing" has been introduced.

Basic concept of intelligent manufacturing is shown in Fig.1. The main components of the system are knowledge, sensors and actuators. Actuators are driven by the information generated through knowldge and outputs from sensors.

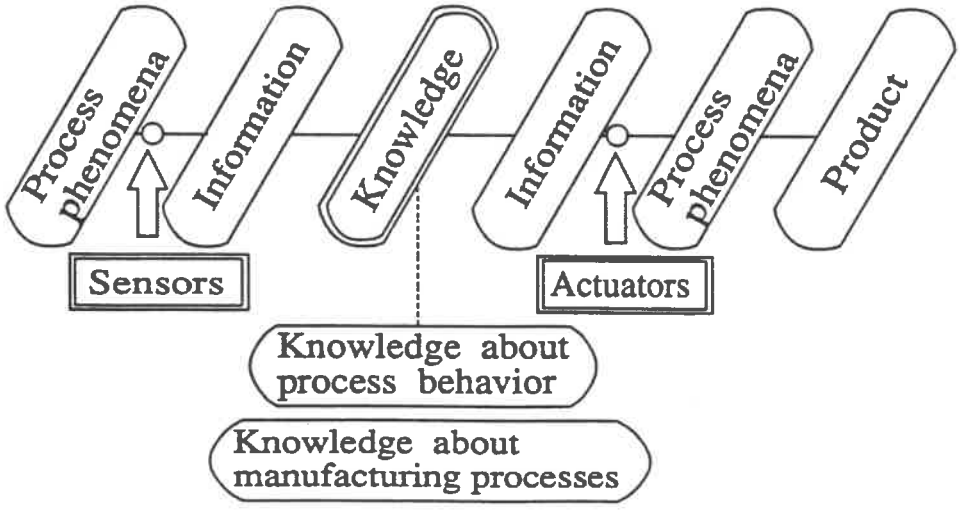


Fig.1 Basic concept of Intelligent Manufacturing

On the other hand, many kinds of physical phenomena occur in manufacturing and many potential types of information are also produced in manufacturing, as shown in Fig.2.

A fundamental block diagram for intelligent manufacturing system, which is based on above concept, is shown in Fig.3.

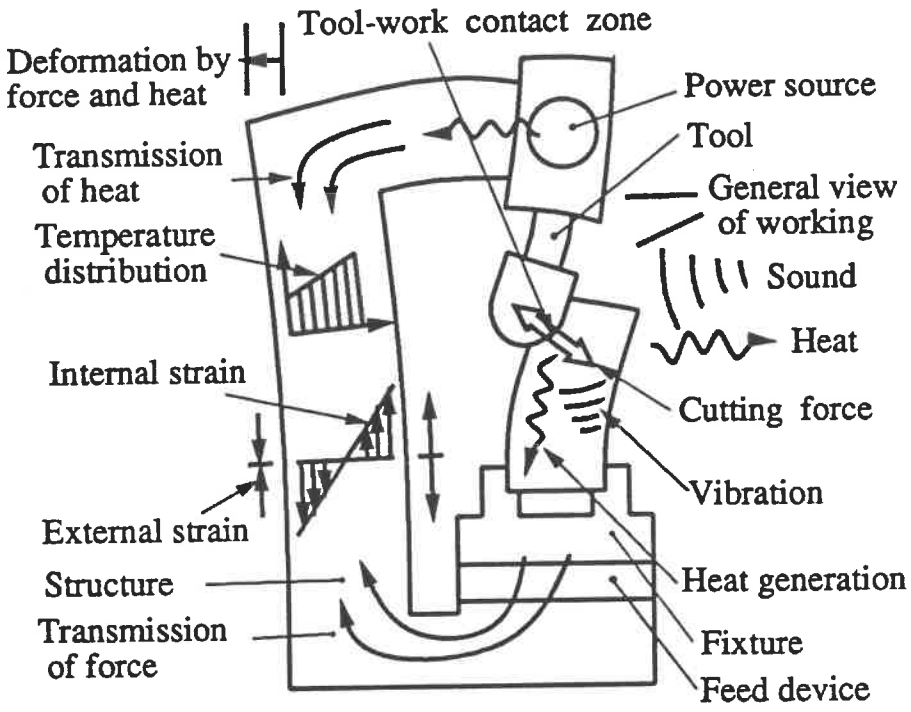


Fig. 2 Fundamental physical phenomena in machining and potential types of information produced

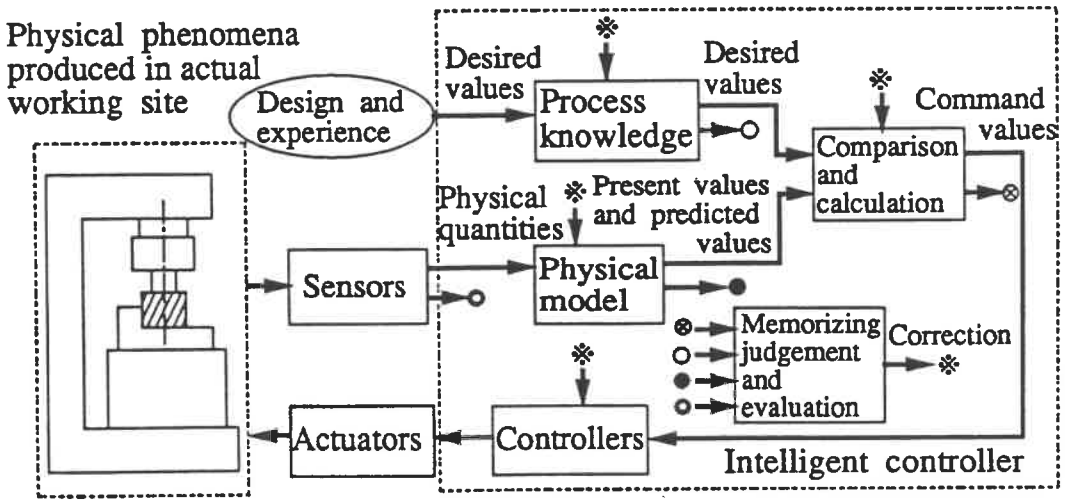


Fig. 3 Fundamental structure of intelligent manufacturing

Process phenomena are modeled by an independent analysis, and the relationships among the dominant parameters are clarified in advance. Information is extracted by sensors, and is input to the physical model. As a result, the current state of the process is

clarified and future state can be predicted. These predicted values are compared with the desired values which are determined from the design specifications. When any discrepancy exists, the parameters are modified in order to satisfy the design values, and the required motions of the actuators are calculated and sent to the controller to drive the actuators. This kind of system not only has high precision and reliability but is also very familiar for human beings and executes the various requirements from the human being honestly and faithfully.

The following three kinds of hardware are indispensable to the intelligent manufacturing.

(1)Sensors to take in various informations before, during and after processing, such as multi-axis force sensors,

(2)actuators to convert the information which is produced through knowledge into actual movement, such as fine positioning mechanism and active structures,

(3)deformation elements to allow for fail-safe or increased compliance behavior, such as fail-safe system.

An example of hardware inplementation of the concept of intelligent manufacturing as applied to a machining center is illustrated in Fig.4.

By using the above concept and developed elements, many kinds of applications such as precise cutting by lathe (Fig.5), prediction system of drill breakage and automatic grinding system are tried . Moreover, nanometer manufacturing robot system under stereo-view SEM, active controlled machining center and tele-operation of manufacturing system using artificial satellite are tried. And the validity of the concept is confirmed.

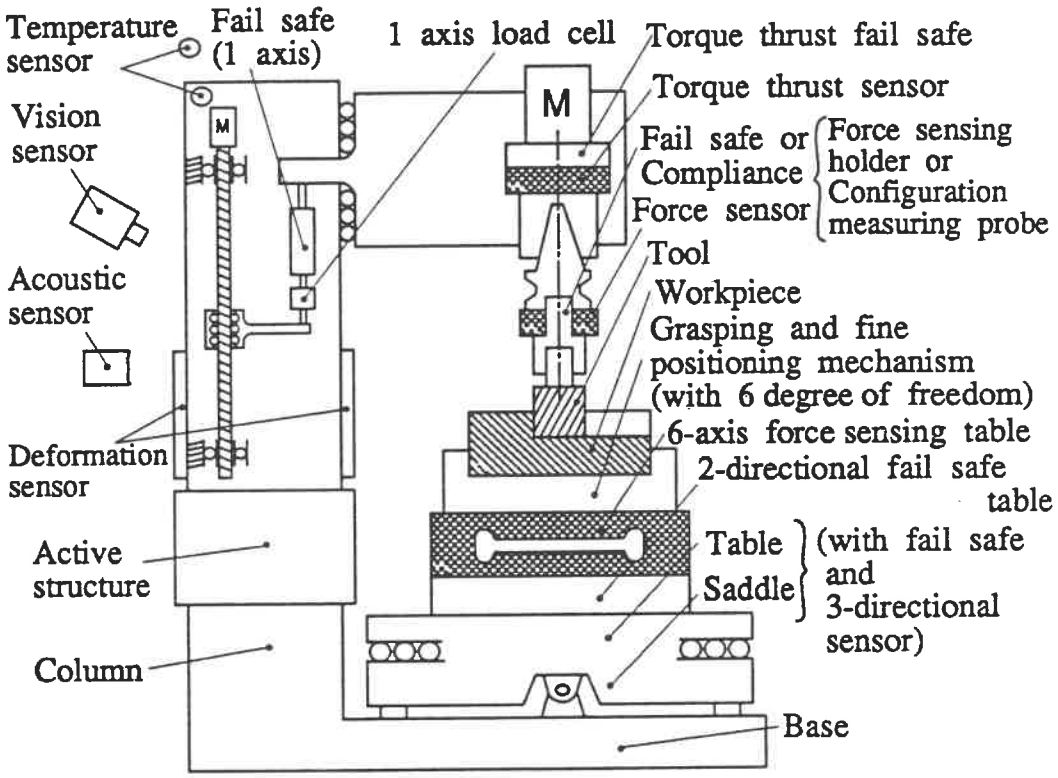


Fig. 4 Hardware structure of machining center where the concept of intelligent manufacturing is applied

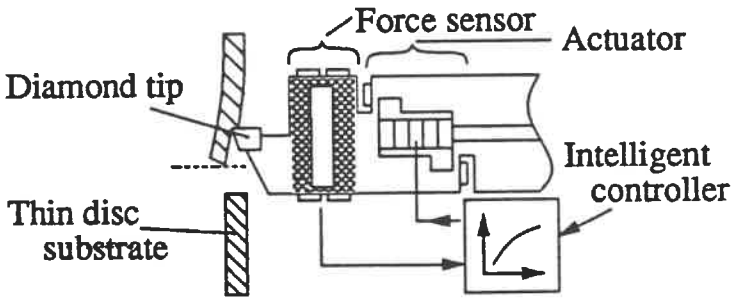


Fig.5 An application example of intelligent manufacturing

Through the intelligent manufacturing system, manufacturing technology will enter a new stage of development. Moreover, this system will open new possibilities for manufacturing in "alian" worlds such as on the atomic scale or in space.

Statistical Concept of System Functional Tolerancing Including Operational Conditions

Abstract

Valentin M. Izraelev

Storage Media Laboratory
3M Company
St.Paul, MN

Background

The system functional quality depends on two major factors - the manufacturing (fabrication) quality and the condition under which the system is used. Manufacturing quality is regulated by imposing tolerances on the errors of fabricated components and assembling. Statistical Methods for the system manufacturing tolerancing are known and well established. Usually the system manufacturing error is represented by parameters of the normal distribution.

Presently the errors induced by operational conditions are represented in the system error by their minimum- maximum. That leads to overestimating the system functional error and increasing production cost. An alternative to that is a statistical representation of "conditional" errors which comes closer to "real-life" situations.

Concept

Several models for statistical representation of operational "conditional" errors with different probabilities allocated to the minimum - maximum limits have been developed. Those probabilities range from 50% in Model II to uniform assignment on the scale between minimum and maximum in Model V.

After choosing the proper models for "conditional" errors and knowing

about manufacturing error, the statistical parameters of the system error distribution is determined by using Method of System's Moments or by Direct Method.

First four moments of the system error allow to approximate the system p.d.f. by one of Karl Pearson's curves or to find percentiles of approximated distribution. For the given risk a the lower and upper limits of the system error are found. If we choose the given risk a as an upper boundary, the real risk associated with the choice of Models I-V for "conditional" errors will diminish with the number of Model. It means that each Model gives an overestimated system functional error with the biggest and the smallest overestimates for Model I (present approach) and Model V respectively.
There are a lot of situations such as interchangeability (for instance in tape drives) when the main interest is the differences between errors. Statistically the interchangeability process can be viewed as a sampling from the population of the system errors with the number of observations in each sample equal two. The exact expression for the distribution of the sample range has been known only for the sampling from Uniform and Normal original distributions. And they can be used as lower and upper boundary on the system functional error range probability function. That approach gives additional relaxation for manufacturing tolerances.

Conclusion

This work is an attempt to introduce a new concept of treating statistically the system functional tolerancing including errors induced by operational conditions which are represented presently by their limits. Several models with different probability allocated to the limits are developed for statistical representation of the operational condition errors. Simple statistical methods are employed for the practical solution of the system error evaluation.

These methods offer opportunities for the system optimization, the manufacturing tolerances relaxation, and therefore they have a lot of potentials for cost reduction. The manufacturing tolerances may be increased up to several hundred percents.

Statistical functional tolerancing is being considered to be tested on two advanced fixed head drives under development at 3M.

Microtilt Controlled Chuck Faceplate for Ultra-precise Single-point Diamond Machining

A E Gee[1], C-L Chao[2], J Duduch[1,3], H Mizuno[1,4], Y Hara[1,5], H Aoyama[1,6], K Flynn[1], K E Puttick[7], M Rudman[7], S G McCandlish[8].

1 School of Industrial and Manufacturing Science, Bldg 40, Cranfield Institute of Technology, Cranfield, Bedford, MK43 0AL, UK
2 Chung-Cheng Institute of Technology, Ta-Hsi, Tao-Yuan, Taiwan, R.O.C. (formerly at 1.)
3 On leave from Dept Engenharia Mecanica, Univ Sao Paulo, Sao Carlos, SP, 13560, Brazil
4 On leave from Nippon Seiko K.K., 1-5-50, Kugenumashinmei, Fujisawa, Kanagawa, 251 Japan
5 On leave from Kobe Steel Ltd, 1-5-1, Takatsukadai, Nishi-ku, Kobe, Hyogo, 651-22 Japan
6 Precision Engineering Dept, Shizuoka University, Faculty of Engineering, 5-1, Johoku 3 Chome, Hamamatsu-city, Shizuoka 432 Japan (on sabbatical leave with SIMS, Cranfield, July 1990 - February 1991)
7 Dept of Physics, University of Surrey, Guildford, Surrey, GU2 5XH, UK
8 School of Engineering, University of Exeter, Exeter, EX4 4QF, UK

(this work, at Cranfield Institute of Technology, has benefitted from contributions from a number of sabbatical visitors, graduate students and collaborators).

1 Introduction: Philosophy and Requirements

Recent requirements and developments in machining indicate that it may be necessary to align and position tools and surfaces in-process on machine-tools and spindles, to sub-micrometre levels; (for example see Refs.[1,2]). In the case of surfaces, to check and/or control position and alignment requires compatible metrology with sub-micrometre capability. Optical interferometry is well-matched to this and has traditionally been used when the test-piece is stationary; as part of earlier work [3], the implications of dynamic interferometric monitoring of a rotating test-piece have been addressed.

This paper addresses mechansm(s) and associated measurements required to make automatic adjustments to an interferometrically-monitored test-piece mounted on a facing lathe spindle.

2 Design-premises

On a facing lathe, the workpiece is required to execute single-degree-of-freedom rotation whilst either the spindle or tool is cross-fed in a plane orthogonal to the spindle. Other motions are required to be maintained at zero (although, for facing, ortho-axial rectilinear motions cause only second order effects). The three degrees of freedom which have principal effect on the shape of the machined surface are rectilinear axial motion and tilts about two axes orthogonal to that of the spindle.

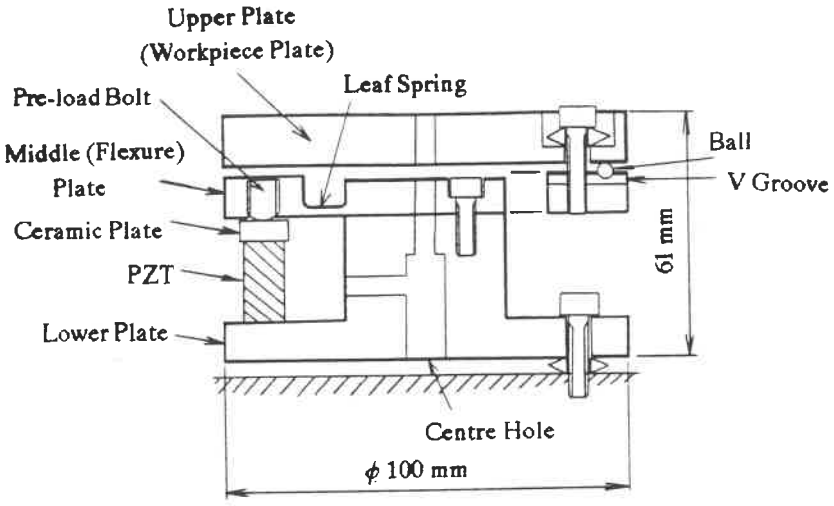


Fig 1 Plan of face-plate tilt-stage

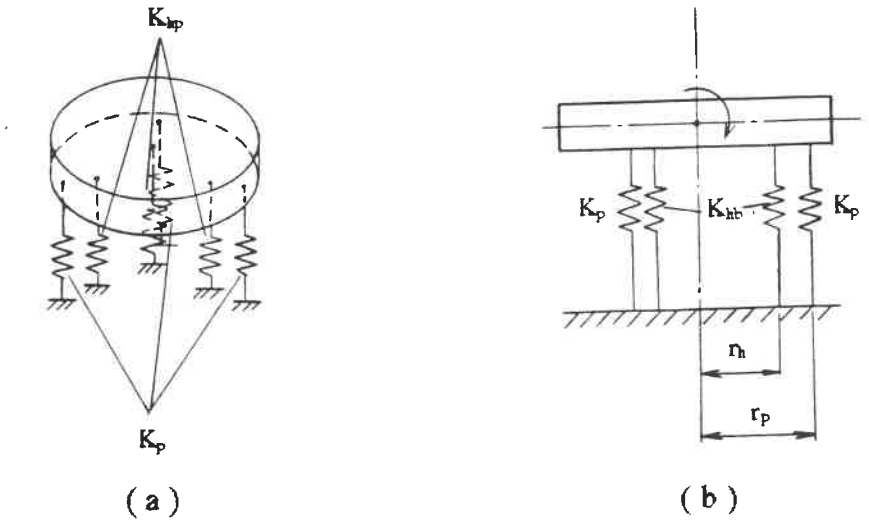


Fig. 2 Models for the analysis of (a) axial stiffness (b) tilt stiffness.

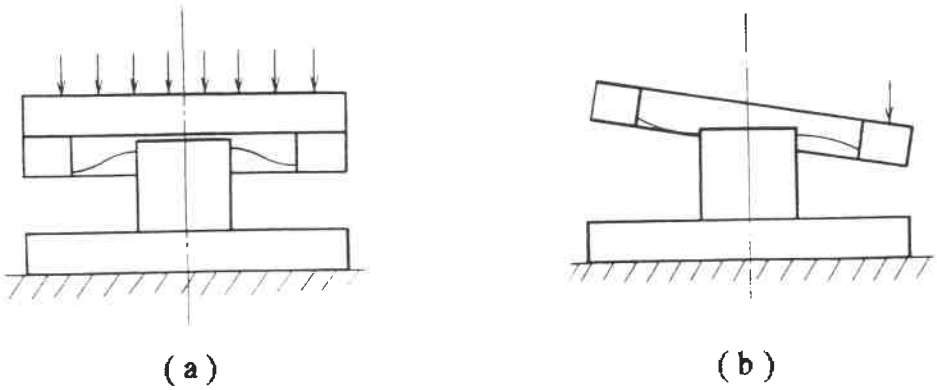


Fig. 3 Action of leaf-spring (a) Parallel mode (b) Bending mode.

IMPROVEMENT OF THE PRECISION OF MACHINING BY THE USE OF INERTIAL AND GRAVITATIONAL DRIVES

E.S. Geskin, M.A. Shimanovich
The Center for Manufacturing Systems and Department of Mechanical and Industrial Engineering
New Jersey Institute of Technology
Newark, NJ

Motors and transmissions of machine tools, especially of precision machine tools not only deliver needed energy for the cutting process, but also conduct necessary information in order to obtain the specified shapes, dimensions and tolerances of the product. Energy transmissions and transformations in the drives are inevitably accompanied by disturbances and, accordingly, by distortions of this information. These effects are especially pronounced for high accuracy machine tools, as well as for more sophisticated machines.

The most critical cases are ultra-high-precision machine tools with contactless spindle bearings and slideways (hydrostatic, gasostatic), where the errors of machined surface are measured in nanometers (nm). The errors introduced by the competently designed contactless bearings and slideways proper are practically equal to zero. With some reservations, it can be stated that accuracy of such state-of-the-art bearings and slideways is unlimited. For example, repeatability of rotation trajectory of the spindle of a diamond turning machine is less than 1 nm, and the shape accuracy of the trajectory itself can be 2-3 orders of magnitude better than accuracy of the bearing surface shape.

Therefore, the machining accuracy is determined in many respects by disturbances not of spindle bearings and support slideways, but by disturbances of drives and transmissions of both spindle and carriages (motors, belt and gear transmissions, couplings, lead screw-nut transmissions, etc). The higher the nominal or realizable power of motor or transmission, the higher the level of those disturbances and their influence on the accuracy of machining.

This presentation is based on an extensive experience in development and adjustment of ultra-high-precision diamond turning machines for finish machining of Winchester drive discs, laser metal optic, of diamond milling machines for finish machining of ultrasonic optic from single crystals, etc. During these programs, various spindle drives have been developed and tested, such as precision belt drive from motor to spindle; rubber-pin coupling; elastic, contactless hydro- and gasostatic and magnetic couplings; electric spindles; two-drive spindle drive, containing heavy-duty drive for setting in motion, braking of spindle and for rough machining and low-power drive for finish machining. Also, very careful vibration protection was needed, such as vibroisolation of the machine and low vibration level of the surroundings.

Each new drive development allowed to improve shape accuracy and reduce roughness of the machined surface. However, there was a ceiling of shape accuracy and roughness which was impossible to overcome (non-roundness R = 40 nm and roughness Rz = 50 nm). Each transmission, coupling, motor design had their own disturbances and influenced in one way or another accuracy of the machined surface.

The best results was obtained from using inertia drive for the spindle when the finish machining was performed by utilizing the accumulated kinetic energy of the spindle during the free run-out of the spindle. The acceleration, braking of the spindle, rough machining are carried out while the spindle is driven by the motor. However, the finish machining of the workpiece is carried out with the motor turned off and disconnected from the spindle supplies of kinetic energy. Since the energy is not supplied to the spindle, the motor and the

The specification called for alignment of a prepared specimen prior to performing final cuts, requiring adjustment of tilts (and possibly axial position).

3 Optical Interferometry to Monitor a Rotating Test-piece

The requirement to monitor and adjust the alignment of a chuck reference-face and/or work-piece, equates to that of aligning an interferometer whilst the test piece is rotated about an axis whose alignment coincides with, or is close to, that of the normal to the interferometer reference-surface. In order to be able to use interferometry dynamically on a rotating machine it is necessary to understand the expected set of fringe-dynamics; this has been the subject of preceding work [3,4].

4. Controlled Tilt-stage

4.1 General

In order to provide for adjustment of tilts about axes orthogonal to that of the spindle, an 'active' face-plate/stage was required. Essentially the choice was between a flexure stage pivotted about two (orthogonal) axes intersecting on the spindle axis, and one pivotted about three (offset) axes mutually disposed at 120° arc. It was intended that piezoelectric actuation should be used, two and three actuators being needed respectively for the arrangements considered.

Because the two-axis system left an assymetrically disposed cantilever (opposite each actuator), it was rejected in favour of a three-axis/three-actuator arrangement which provided a symmetrical closed structure. In this it was accepted that motion of any one actuator generates tilt about the other two and, perforce, 'parasitic' axial displacement of the centre of the stage/test-piece. In order to achieve tilt about an axis passing through the spindle-axis it was accepted

4.2 Face-plate Tilt-stage

Figure 1 shows a sectional diagram of the three-axis arrangement orientated for use on a vertical axis machine. It consists of three separate overlaid circular plates and three piezoelectric stack actuators (NEC Corp. 10x10x20mm, with displacement of 15um corresponding to the maximum applied voltage of 150V). The top (duralumin) plate acts as workpiece holder. It is mounted via a three-ball/three-groove Kelvin clamp on the middle (steel) plate which is relieved around an annulus to form a flexure of three symmetrically arranged radial leaf-springs supporting an outer ring on a central hub. The hub is mounted on an axial pillar, part of the lower plate which is turn fixed to the machine-spindle.

The three piezoelectric actuators are placed on the leaf-spring radials (extended). Each controls the displacement of the outer ring of the flexure. A screw-adjustable preload is incorporated to avoid backlash and ceramic plates are inserted between the preload screws and the actuators to act as buffers to protect the actuators. The actuators are arranged to be driven from external signals via slip-rings on the spindle.

4.3 Analysis

The design is based upon a nominal response bandwidth specification of 500Hz which translates to requirements for tilt and axial stiffness. A simplified model is shown in Figure 2 and the analysis based upon the following:

(i) the three leaf-springs and the actuators acting as parallel pairs of springs of stiffness K_h and K_p respectively;

(ii) the tilt- and axial- mode resonant frequencies are determined by the stiffness of the leaf-springs and the mass of the moving parts (ie the outer-ring of the flexure and the upper plate plus any work-piece);

(iii) the leaf-springs behave in a parallel mode (where the stiffness is K_{hp}) for axial loads and in bending (where the stiffness is K_{hb}) for tilt-moments as shown in Figures 3(a) and (b) respectively.

5 Performance of the Tilt-stage

The characteristics and performance of the stage were as follows:

Range:	200urad
Axial Stiffness:	180N/um
Tilt Stiffness:	0.11Nm/urad
Bandwidth:	480Hz

6 Summary

A micro-tilt stage has been designed and developed which is able to align a workpiece to super-precise levels on a spindle, where necessary, and whose performance is adequate for single-point diamond-turning to sub-micrometre levels.

7 References

1 S R Patterson and E B Magrab, "Design and Testing of a Fast Tool Servo for Diamond Turning", Precision Engg., vol 7 pp 123-128 (1985)

2 S Moriyama, "Precision Aspheric Generators for Off-axis X-ray Mirrors Asymmetric Aspheric Lenses" pp 242-54 of Seyfried et al (eds) "Progress in Precision Engineering" Springer, Berlin, 1991 (Proc.IPES6, Braunschweig, Germany, May 1991)

3 A E Gee, S G McCandlish and K E Puttick, "Interferometric Monitoring of Spindle and Workpiece on an Ultra-precision Single-point Diamond Facing Machine" Proc SPIE vol 1015 pp 74-80 (1988)

4 A E Gee, "Interferometric Form Metrology Applied to a Rotating Surface", in J C Dainty (ed) 'Applied Optics Digest' (Proc. Optics 90 Conference, Nottingham, September 1990), Inst.P, Bristol, 1990, pp 109-110

couplings (transmissions) are not rotating and their disturbances are absent. The spindle rotates by itself in its gasostatic bearings, and its kinetic energy is spent to overcome gas friction in the bearings, friction between the spindle and the workpiece and the ambient air, and the cutting forces. Use of this technique for turning of Winchester discs resulted to roughness up to Rz = 10 nm and in case of turning of laser metal optic components - Rz = 4-5 nm.

The run-out time of the spindle was 30 min in the machine for turning Winchester discs and 150 min in the machine for turning laser metal optic components. The spindle RPM in gasostatic bearings were varied, practically, as an exponent. This effect was used for cutting speed stabilization for turning of face surfaces of the discs.

The friction torque and inertia about the spindle axes was so selected, that the cutting speed in the beginning and the end of the cutting pass in the direction from the disc center outside was the same. At present, the analytical functions allowing to calculate the run-out of the spindle with accuracy up to 5% are available. The energy consumption by the cutting process (diamond turning) is about 5% from the energy loss on friction in the bearing and in the ambient air [1].

The feed drive in the described machines was using the outside energy, and thus generated certain disturbances. The minimum level of disturbances and therefore the maximum accuracy of machining can be achieved using a machine, on which the process of finish machining of the product is carried out while its motors and transmissions are disconnected, using only the accumulated kinetic energy of the spindle and potential (gravitational) energy of the tool support. Moreover, the energy necessary for the power supply of the contactless spindle bearings and tool support slideways are also provided from supplies of accumulated energy. Such a machine tool would provide ultimate low roughness, determined only by the structure of the machined material, namely by the size of atoms, that is for copper and aluminium is about 0.3 nm.
At present, there are operational dozens of ultra-high-precision machine tools with inertia spindle drives for diamond turning of Winchester discs and laser metal optic components. The next step is development of machine tools for finish machining, in which motors and transmissions for feed drives are also eliminated, and the feed motion is carried out by the accumulated potential energy of the tool support. Of course, use of accumulated energy by work organs allows to eliminate only disturbances from the motors and transmissions. Influence of other perturbing factors, e.g. ambient vibrations, remains.

The machine tools with inertia spindle drive for finish machining of Winchester discs can be made with a device for replacement of workpieces without stopping the rotating spindle. There is no spindle motor in this machine, and the spindle will receive a supply of kinetic energy from the motor of device for replacement of workpieces during replacement, placed outside the machine tool.

REFERENCES

1. M.A. Shimanovich, E.M. Zakharevich. Aerodynamic Frictional Losses of a Spindle in an Ambient Medium. Soviet Engineering Research, No. 7, 1985, pp.71-73.

PRECISION GRINDING OF THIN QUARTZ WAFERS

Joseph Hosler and Thomas Bifano
Aerospace and Mechanical Engineering
Boston University
Boston, MA 02215
(617) 353-5619

Introduction

Quartz wafers are ideally suited for use as high-frequency bulk acoustic wave mechanical resonators. Their piezoelectric properties permit an efficient tranfer from electrical energy to mechanical energy. The frequency and quality of mechanical resonance, however, are determined by the geometry and surface quality of the wafer: very thin wafers of uniform thickness are needed to obtain high frequency resonance with minimal frequency broadening.

Limitations of Present Production Method

The present method of wafer production, grinding to shape and then lapping to remove the fractured surface produced by the grinding, carries many limitations. Even though grinding is an inherently accurate contouring process, it usually entails material removal through a brittle fracture mechanism producing a poor surface finish. The damaged surface created by the grinding must then be removed through a ductile process, such as polishing and/or lapping. As this step removes the fracture surface, it also destroys the contour produced by the grinding, creating a larger thickness variation than desired. The limit of thickness tolerances achievable in double-sided lapping of thin wafers is currently 1 μm. These limitations on thickness tolerance directly affect the frequency quality of quartz resonators that can currently be produced.

New Quartz Wafer Production Method

This paper describes an improved method for manufacturing thin quartz wafers. By integrating three current technologies of precision engineering (ductile-regime grinding, low-stress vacuum fixturing, and real-time sensing and actuation), we are able to guarantee uniformity of the wafer thickness while perserving a good surface finish. We have constructed a two axis ultraprecision grinding apparatus at Boston University which uses these three technologies to produce wafers as thin as 25 μm and with thickness variation of less than 0.4 μm.

Ductile-Regime Grinding Through recent advances in precision engineering, it has become possible to grind extremely brittle materials in a ductile material-removal

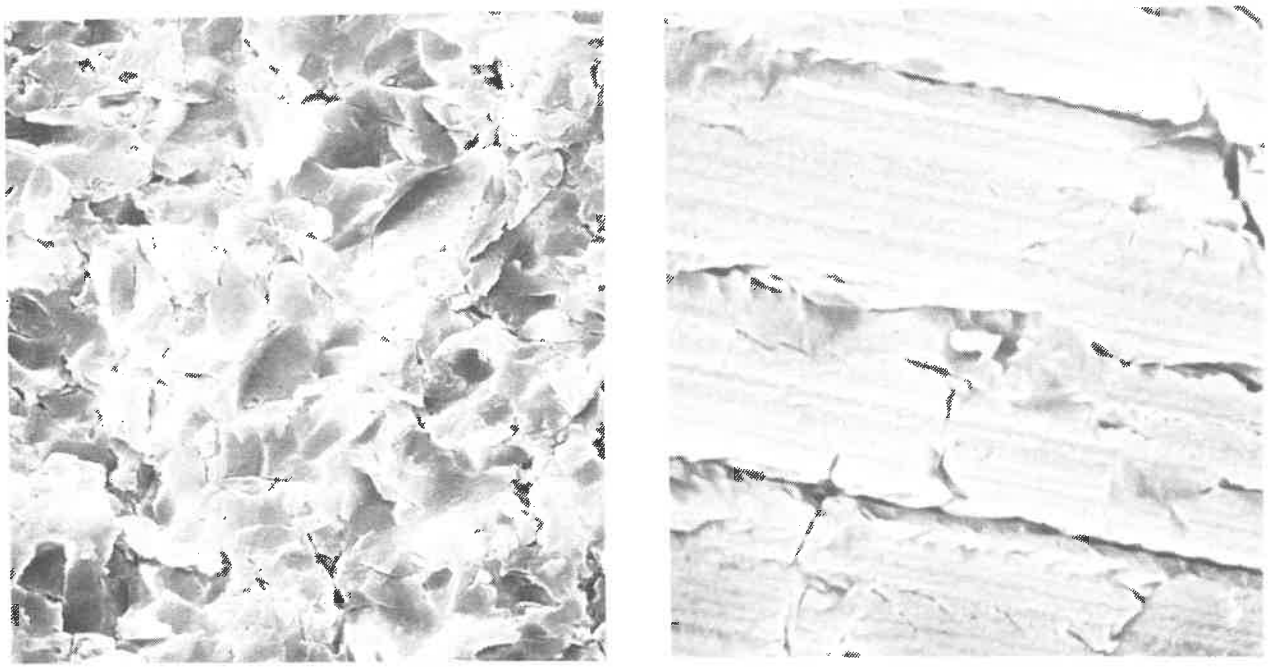


Figure 1. SEM photographs of crystalline quartz (6000x). Left: prior to grinding on precision grinding apparatus. Right: Ground on precision grinding apparatus after 20 μm of infeed in 500 nm steps. Crossfeed was approximately 2.5 μm/rev.

regime, with almost no fracture of the material surface. This new technology is called *ductile-regime grinding,* and it provides the ability to obtain accurate contours and mirror-quality finishes in a single operation. The first reported research on ductile-regime grinding in the United States was performed by Bifano [1]. In this study, ductile-regime grinding was analyzed on both experimental and theoretical grounds. A theoretical model was developed for the transition from brittle-regime grinding to ductile-regime grinding, predicting that the brittle-to-ductile transition is primarily a function of the abrasive-grain depth-of-cut. That is, reducing the depth-of-cut or infeed rate of the grinding wheel below a certain critical value will ensure that the material-removal mechanism is one of plastic flow, resulting in an optical quality surface. This eliminates the need to lap the surface, leaving the contour, produced by grinding, intact. In our application, we have used infeed rates up to 500 nm and crossfeed rates of 2.5 μm/rev and have been able to get an appreciable amount of ductility in crystalline quartz (see Fig. 1).

Wafer Fixturing Low-stress vacuum fixturing was chosen as the method to secure the thin wafers during machining. The objective was to fasten the wafers in such a way that the deformations caused by the fixturing are insignificant with respect to the thickness tolerances required for the wafer. A vacuum actuated porous ceramic chuck was chosen for its ability to provide a uniform holding force over the entire surface of the wafer. The porous ceramic gives us the capability to create a large number of vacuum ports, each port having a very small diameter. In our case, we chose a pore diameter of approximately 25 μm. By using this pore size, the fixturing forces are distributed thinly enough over the wafer surface so that the deflection over any individual port is infinitesimal compared to the thickness of the wafer. By modelling the

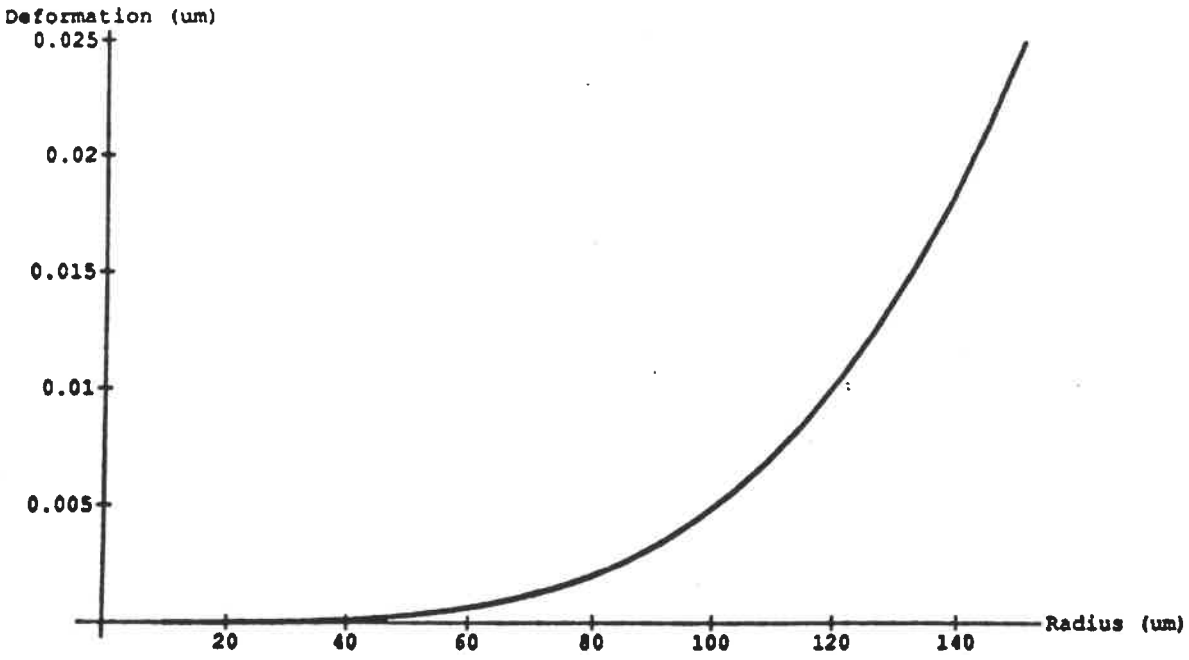


Figure 2. Deflection of the center of a circular plate simply supported at its edges with a uniform pressure of 75 kPa acting over the entire top surface.

ceramic chuck pores as simple circular supports for the wafer, and assuming a uniform pressure of 75kPa, the deflection errors due to chucking can be estimated [2] (see Fig.2). It is found that with pore sizes of 25 μm, wafers that are 25 μm thick will see less than 1 nm of deflection. Since the flatness of the machined wafer will only be as good as the flatness of the vacuum chuck, it is necessary to machine the surface of the vacuum chuck in place on the same apparatus that will be used to machine the wafers. We have produced a chuck flat to less than 150 nm (see Fig. 3). With this flatness, we may grind each side of the wafer to a flatness well above the limits of our machine while keeping the thickness tolerances desired.

The combination of ductile-regime grinding and low-stress vacuum fixturing enables us to produce flat, damage-free wafers of uniform thickness in a single operation, reducing the amount of processing needed. Because this grinding process utilizes real-time, closed-loop position feedback, the production of quartz wafers of uniform thickness becomes a reliable and accurate process.

Precision Grinding Apparatus

A two axis grinding apparatus has been constructed at Boston University to serve as a test bed for the experimental investigations. This apparatus incorporates two linear air bearing slideways and a rotary air bearing spindle into a structure with an overall stiffness of 50 MN/m. In our setup, we have mounted the workpiece on one of the linear slideways, which is used as a crossfeed mechanism. To maintain a grain depth-of-cut required to obtain ductile material removal, both the crossfeed axis and the

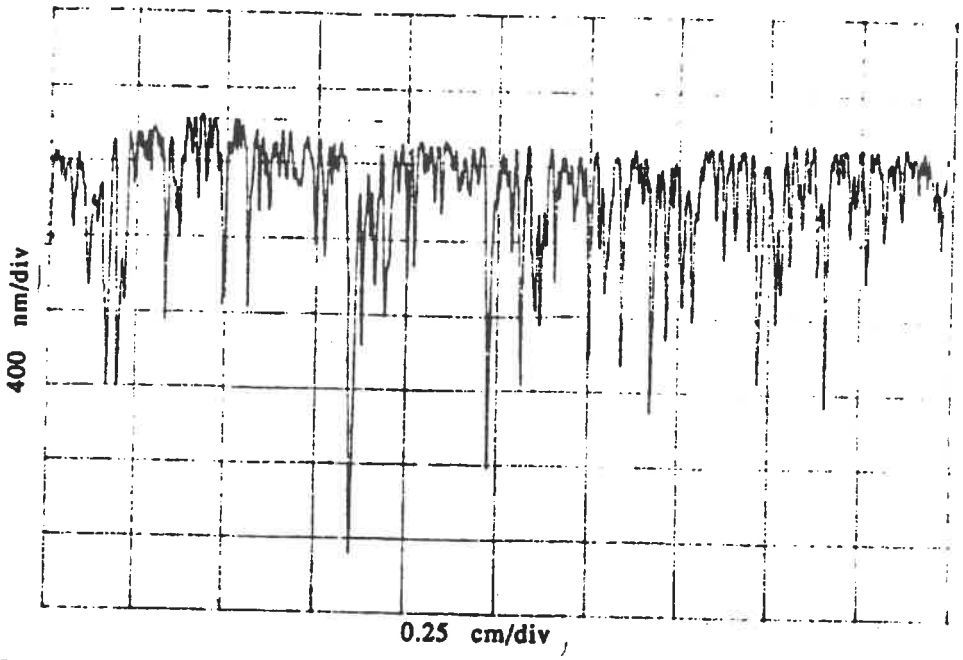


Figure 3. Profile of the porous ceramic chuck ground on the precision grinding apparatus. P-V height is approximately 150 nm.

infeed axis are controlled in real time with a precision of ~25 nm. For the crossfeed, this is accomplished using a piezoelectric "inchworm" actuator, which has a resolution of 5 nm and a range of 50 mm. Infeed of the workpiece into the grinding wheel is accomplished using a fast piezoelectric actuator, controlled in real time by closed-loop feedback from a capacitance gage reading off of a reference flat, which has a flatness of better than 25 nm per inch. The reference flat is mounted behind the INVAR workpiece holder, reducing the Abbe offset and thermal growth errors. The grinding wheel is cup shaped, with all grinding ocurring on the rim of the cup. The wheel is mounted on a rotary air bearing that is driven by a DC servomotor.

Conclusion

We have shown that 25 μm thick quartz wafers with a thickness variation of 0.4 μm can be produced consistently through a ductile grinding process with low-stress vacuum fixturing and real time control. This extends the range that bulk acoustic wave quartz resonators may operate and provides a quality better than has been achieved by the current manufacturing processes.

References

[1] Bifano, T. G., "Ductile-Regime Grinding of Brittle Materials." Doctoral Dissertation, North Carolina State University, Raleigh, NC, 1987.

[2] Roark, R. J., and Young, W. C., "Formulas for Stress and Strain." 5th ed., McGraw-Hill Book Company, 1982.

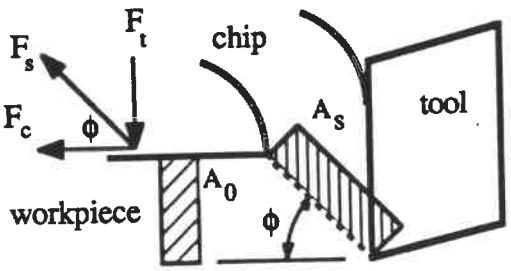


Figure 8: Forces on shear plane

The dynamic shear stress, or flow stress, has been shown to be a material dependant constant for wide ranges of cutting parameters. Using the nomenclature of Figure 2, the flow stress can be written in terms of only forces acting on the rake face.

$$\tau_s = \frac{F_s}{A_s} = \frac{F_{Y_r}\cos\phi - F_{Z_r}\sin\phi}{A_0 / \sin\phi} \qquad (5)$$

which can be simplified using the relationships $A_0 = Ld$ and $F_{Zr} = \mu F_{Yr}$.

$$\tau_s = \frac{F_{Y_r}}{L\,d}\left[\sin\phi\cos\phi - \mu\sin^2\phi\right] \qquad (6)$$

In terms of the model parameter, C_1, where $C_1 L\,d = F_{Yr}$:

$$C_1 = \frac{\tau_s}{\sin\phi\cos\phi - \mu\sin^2\phi} \qquad (7)$$

C_1 can be expected to stay constant if the geometrical factor, $1/(\sin\phi\cos\phi - \mu\sin^2\phi)$ does not change. This depends solely on the shear angle if the friction coefficient remains constant. The theory of Merchant predicts a shear angle of 40° using $\phi = 45 - \beta/2 + \gamma/2$ where:

β = friction angle = $\tan^{-1}(\mu)$ (11° if $\mu = 0.2$)

γ = tool rake angle (0° used here)

Chips from the cutting tests were collected and measured in the SEM to calculate ϕ from the ratio of cutting depth to chip thickness. If these are equal, the angle is 45 degrees. The uncut chip thicknesses (cutting depths) were 0.3, 0.6, and 1.2 μm. Examinations at 20 000X yielded chip thicknesses from which shear angles of 39°, 41°, and 39° were calculated respectively. A variation of shear angle from 30° to 50° in equation (7) only produces a variation from 2.4 to 2.7 in the geometrical factor relating C_1 to τ_s. Therefore, it is not surprising that C_1 was found to be constant for depth of cut and tool wear. C_1 can be considered as a material property related to flow stress by a constant.

Conclusions

Tool force measurement can lead to understanding of the diamond turning process which can be applied to improving the product quality. For example, force information might enable positioning correction to be made based on cutting loop stiffness to maintain form accuracy. The force modeling can also be used with experimental data to monitor tool wear quantitatively. Knowledge of how tool wear affects surface finish can lead to improvements in finish.

References

[1] J. T. Carroll, T. A. Dow, J. S. Strenkowski; *"Tool Force Measurement and Prediction in Diamond Turning"*, SPIE Vol 676, Ultraprec.Machng. & Automtd. Fabrctn. of Optics, 1986

[2] J. T. Black; *"Flow Stress Model in Metal Cutting"*, Journal of Engineering for Industry, Vol.101,pp403-15; Nov.1979

[3] M. E. Merchant; *"Mechanics of the Metal Cutting Process. I. Orthogonal Cutting and a Type 2 Chip,"* Journal of Applied Physics, Vol.16,pp267-75; 1945

SHORT-CRACK TOUGHNESS DETERMINATION FOR BRITTLE MATERIALS

Stanley M. Smith and Ronald O. Scattergood
Precision Engineering Center
Raleigh, NC

INTRODUCTION

Determination of the short-crack fracture toughness is important for prediction of the critical depth of cut at which the mode of material removal during precision machining undergoes the brittle to ductile transition. Indentation techniques are well suited for determination of the short-crack toughness. *In situ* measurement of crack size as a function of applied stress allows determination of the fracture toughness. Analysis of the data obtained for soda-lime glass showed that stress-intensity factors based on the classic half-penny crack shapes overestimate toughness values and produce an apparent increase of toughness with crack length. This is due to a constraint on crack shape imposed by primary lateral cracks. Models based on elliptical cracks have been developed to account for the crack shape effects.

BACKGROUND - FRACTURE MECHANICS

The radial/median indentation-crack system (radial) is treated in most analyses as having a half-penny configuration [1,2]. In this section, the fracture mechanics relations for half-penny indentation cracks under bending stresses are derived. The crack shapes can be semicircular, but this need not be the case. Appropriate modifications for departure from half-penny geometry will be made later.

For a half-penny radial crack generated by Vickers indentation, the stress intensity factor K for a crack length c, indentation load P and applied stress σ is given by [2]

$$K(c) = K_{bend} + K_{residual} = \psi \sigma c^{1/2} + \frac{\chi P}{c^{3/2}} \quad (1)$$

The K_{bend} term corresponds to the stress intensity factor for an applied bend stress σ while the $K_{residual}$ term results from the indentation residual-stress field due to an indentation load P. ψ is a crack-shape factor that can be calculated for given crack geometry. χ is a residual-stress factor which cannot be determined *a priori*, but must be calibrated from experiments.

Fracture toughness was determined by *in situ* measurement of crack length as a function of applied stress, ie., the stable growth of an indentation precrack is monitored using a bending stage on an optical microscope. The σ *vs* c data obtained establish the relation $\sigma(c)$ for fixed P. If ψ and χ are known, Equation (1) with $K(c) = K_c$ will give the toughness K_c directly for the available range of crack lengths. Alternatively, Equation (1) can be rearranged to give

$$\sigma c^2 = \frac{K_c}{\psi} c^{3/2} - \frac{\chi P}{\psi} \qquad (2)$$

A plot of σc^2 *vs* $c^{3/2}$ will be a straight line. In this case, only the crack-shape factor ψ is needed for the analysis. K_c/ψ is obtained as the slope of the plot. The intercept $\chi P/\psi$ allows the residual-stress factor χ to be *independently* evaluated at any fixed P. If toughness is not constant (R-curve effect), then the toughness T(c) is determined from the curvature in the plot.

TOUGHNESS DETERMINATION

The description of the experimental setup has been summarized previously [3]. Vickers indentation precracks were placed onto the center of annealed soda-lime glass disks for indentation loads between 5 and 80 N. σ *vs* c data were acquired in a dry nitrogen atmosphere. Stress was applied in a stepwise fashion, the stress rate being approximately 1 MPa/s with hold periods of about 5 s at 2.5 - 5 MPa intervals. Video' images of the cracks were acquired during the hold times.

Values of K_c and χ *vs* P were derived from the σ *vs* c data by plotting according to Equation (2) using ψ = 1.29 [4]. Analysis based on half-penny crack shapes indicated an apparent T-curve behavior, inconsistent with soda-lime glass. In addition, the values of K_c were much higher than the generally accepted value of 0.75 MPa $\sqrt{m}$.

Sub-surface crack shape profiles were observed by examining the fracture surfaces of indented soda-lime glass bars broken in four-point-bending. The crack shape was observed to be in general elliptical rather than half-penny, with the depth dimension of the indentation crack being constrained by primary lateral cracks. Figure 1 summarizes schematically the actual crack geometry based on a large number of cross-section observations.

Elliptical Crack Model To include crack-shape effects into the analysis, a physically realistic model for the constraint imposed by the primary lateral cracks must be developed. The depth a of the lateral crack will be assumed constant for given P. The crack-shape constraint effect will be subsumed into a one-parameter, elliptic-crack model where the ellipticity is governed by the depth of the lateral component. The K_{bend} and $K_{residual}$ terms in Equation (1) were modified to account for crack ellipticity (eccentricity), e = c/a ≥ 1. Ellipticity corrections for K_{bend} follow from the work of Newman and Raju [4]. These authors used finite-element methods to estimate shape factors ψ for elliptical surface cracks subjected to bend loading. The crack-shape factor can then be written as $\psi = \psi(e) = \psi_o \times f_{bend}$ where ψ_o = 1.29 is the value for half-

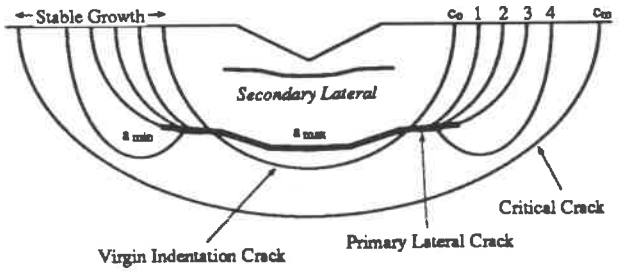


Figure 1: Schematic representation of the observed sub-surface indentation crack geometry.

penny surface cracks. The ellipticity correction is f_{bend} where $f_{bend} = 1$ for half-penny cracks (e=1). A similar crack shape factor $F_{residual} = f_{residual} \times g_{residual}$ for accounting for crack ellipticity in the $K_{residual}$ term was derived using the method of Oore and Burns [5]. $g_{residual}$ is a small free-surface correction factor. $f_{residual} = 1$ for half-penny radial cracks. The calculational procedures employed to evaluate these factors have previously been derived in [6].

Combining the above into Equation (1), the corrected stress intensity factor for the elliptic-crack model becomes

$$K(c) = \psi_o f_{bend}\, \sigma c^{1/2} + F_{residual} \frac{\chi P}{c^{3/2}} \qquad (3)$$

Using Equation (3), the form of Equation (2) appropriate to the elliptic-crack model for constant toughness is

$$\frac{f_{bend}\, \sigma c^2}{F_{residual}} = \frac{K_c c^{3/2}}{\psi_o F_{residual}} - \frac{\chi P}{\psi_o} \qquad (4)$$

K_c and χ can be independently determined by plotting $f_{bend}\, \sigma c^2/F_{residual}$ *vs* $c^{3/2}/F_{residual}$.

To apply Equation (4) to σ *vs* c data, the crack ellipticity e = c/a must be known. Theory asserts that lateral crack depth "a" should scale with the indentation plastic zone size [7], hence, with $P^{1/2}$. The relation for crack ellipticity is then

$$e = \frac{c}{a} = \frac{c}{\lambda P^{1/2}} \qquad (7)$$

where the constant $\lambda = a/P^{1/2}$ provides a convenient means to introduce ellipticity into Equation (4). In this formulation, λ represents the adjustable parameter in the elliptic-crack model. Since the lateral depth varied as shown in Figure 1, minimum and maximum values of the depth were determined at each P using an average of ten measurements in each case. Plots of a *vs* $P^{1/2}$ gave values of $\lambda_{min} = 1.2 \times 10^{-5}$ m/$N^{1/2}$ and $\lambda_{max} = 1.8 \times 10^{-5}$ m/$N^{1/2}$.

Plots according to Equation (4) were linear beyond an initial start-up transient (due to environmental effects). Figure 2 shows the dependence of K_c and χ on P using λ_{max} for radial crack geometry. Both K and χ are nominally constant with P, the average value and standard deviation for all tests being $K_c = 0.54 \pm 0.068$ MPa$\sqrt{m}$ and $\chi = 0.029 \pm 0.007$. Figure 2 also shows values of K_c and χ obtained using Equation (3) for half-penny crack geometry. Application of the ellipticity corrections eliminate the apparent T-curve effect and reduce toughness values to the range expected for soda-lime glass. Kc = 0.54 MPa$\sqrt{m}$ is less than the widely quoted inert-limit value $K_c = 0.75$ MPa$\sqrt{m}$ for soda-lime glass. Choosing $\lambda < \lambda_{max}$ reduces K_c to smaller values, implying that the upper limit λ_{max} gives the best fit using the radial-crack model. Simulations with all of the ellipticity corrections included were carried out to quantitatively estimate the effect of environmental crack growth in dry nitrogen on the results shown in Figure 2 for the best-fit radial crack model [6]. The simulations showed that environmental corrections of 20% increase and 9% decrease must made to K_c and χ, respectively, to get inert-limit values. This gives $K_c = 0.65$ MPa$\sqrt{m}$ and $\chi = 0.026$ as the average inert-limit values for *in situ* tests on soda-lime glass. Given the crack-shape and environmental corrections involved, these are in good agreement with values reported earlier in the literature.

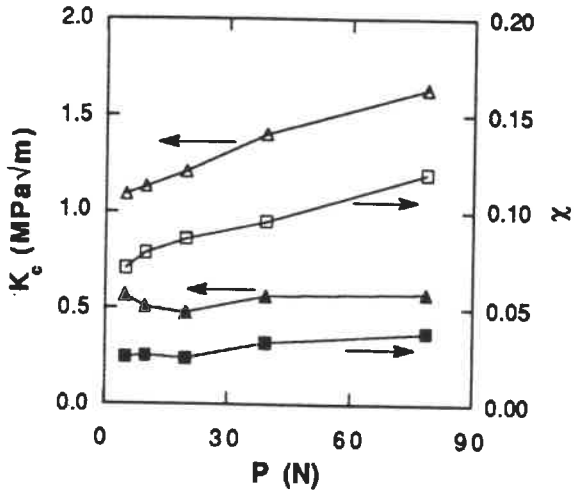


Figure 4: Predictions from the classical half-penny model (open points) and the best-fit elliptic-crack model for radial geometry.

SUMMARY

The *in-situ* method developed here for measuring short-crack toughness has the advantage that K_c and χ can be independently determined. This provides a useful methodology for characterizing short-crack toughness properties of glasses and ceramics used in precision machining applications. However, the present results clearly show that crack shape is a heretofore unrecognized and important factor when using indentation precracks. Shape effects must be taken into account in fracture toughness analysis.

REFERENCES

[1] Antis, G.R.,Chantikul. P., Lawn, B.R., and Marshall, D.B., "A Critical Evaluation of Indentation Techniques for Measuring Fracture Toughness: I, Direct Crack Measurement"., *Journal of the American Ceramic Society*, **64** [9] 533-38 (1981).

[2] Chantikul, P., Antis, G.R., Lawn, B.R., and Marshall, D.B., "A Critical Evaluation of Indentation Techniques for Measuring Fracture Toughness: II, Strength Method", *Journal of the American Ceramic Society*, **64**[9] 539-43 (1981).

[3] Smith, S. and Scattergood, R.O., "Material Factors in Precision Machining", 1990 NCSU *Precision Engineering Center Annual Report.*

[4] Newman, J.C. and Raju, I. S., "An Empirical Stress-Intensity Factor Equation for the Surface Crack", *Engineering Fracture Mechanics*, **15** [1-2] 185-92 (1981).

[5] Oore, M. and Burns, D.J., "Estimation of Stress Intensity Factors for Embedded Irregular Cracks Subjected to Arbitrary Normal Stress Fields", *Transactions of the ASME*, **102** 202-11 (1980).

[6] Smith, S.M. and Scattergood, R.O., "Crack Shape Effects in Fracture Toughness Measurements", submitted to the *Journal of the American Ceramic Society*, 1991.

[7] Cook, R.F. and Pharr. G.M., "Direct Observation and Analysis of Indentation Cracking in Glasses and Ceramics", *Journal of the American Ceramic Society*, **73** [4] 787-817 (1990).

MICROFABRICATION OF ULTRAPRECISION STRUCTURES USING FOCUSED ION BEAM MACHINING

S T Davies

Centre for Microengineering and Metrology
Department of Engineering
University of Warwick
COVENTRY CV4 7AL, UK

Abstract

The application of focused ion beam machining (FIBM) to the production of ultraprecision artifacts is described. A 30 keV ion beam focused to a 0.25 μm minimum probe diameter is used to pattern three-dimensional structures in thin metallic films by direct writing under computer control. Techniques for depth control of the milling process and in-situ metrology are demonstrated. Examples of microfabricated structures with potential applications for micromechanical devices are shown.

Introduction

Patterning and processing of thin film structures are emerging as important application areas of focused ion beam (FIB) technology[1]. Material is removed or deposited on a fine scale by physical sputtering or ion beam induced deposition from the vapour phase respectively. In both cases the spatial resolution achievable is essentially equal to the ion probe diameter employed. Repair of photolithography masks, modification of prototype application specific integrated circuits (ASICs) and IC diagnostics and failure analysis are examples where FIBM is already being commercially exploited. Other areas of current interest include micromachining of high Tc superconductors, micromachining of structures for integrated optoelectronics and x-ray lithography mask repair. Almost invariably, these applications have been limited to scanning the FIB in a simple raster pattern to perform milling or deposition operations. Here, FIBM using digital computer control of beam positioning is described, allowing generation of complex patterns for microfabrication purposes.

Experimental

A schematic of the Elionix 30 keV FIB system used for this work is shown in Figure 1. Ions produced by a liquid metal ion source (LMIS) are extracted and passed through a condenser lens and beam defining aperture. An E X B mass separator allows selection of different ion species and charge states from an alloy LMIS. Ions accelerated down the column are deflected by alignment electrodes and focused by an einzel lens into a 0.25 μm minimum probe diameter in the target plane. The FIB is electrostatically scanned over the target surface. In the applications described below, either singly or doubly

charged gold ions emitted from a gold/silicon alloy source were used.

Software was developed to drive the scan generation electronics under external computer control. Programmable raster scanning was implemented, with variable pixel spacing, line spacing and pixel dwell time. No correction was made for the hardware aspect ratio, so scans programmed with equal pixel numbers in x and y directions were in the ratio 4:3 in terms of physical beam deflection. Algorithms were written to allow image or linescan data capture from a variety of detectors and to display captured data as a function of time or beam spatial position. Routines for generation of vectors, arcs, circles and spirals with variable pixel spacing and dwell time were also implemented.

In-situ calibration of scan field size was done by digitally scanning the beam across a 25 μm gold wire and displaying the summed signal from two secondary electron detectors symmetrically positioned with respect to the target. Signal intensity (arbitrary units) as a function of beam position is plotted in Figure 2(a). Assuming the FWHM of the distribution to be 25 μm, simple scaling gives a scan deflection range of 75 μm. The cursors show the half maxima positions and the 10% to 90% signal intensity changes on the rising and falling edges. The latter corresponds to a displacement of approximately 5 μm which is related to the probe diameter used for the measurement. An advantage of using two independent detectors is that two dimensional surface profile information can also be obtained[2]. Figure 2(b) shows the result of computing the integral of the difference of the two detector signals, and represents the profile of the 25 μm diameter wire. Beyond about ± 70° from the vertical, geometrical shadowing of the detector line-of-sight means that an accurate profile cannot be obtained. The horizontal axis is identical to Figure 2(a) and the vertical axis may be calibrated from the assumed value of wire radius.

Micromachining of thin films

Many current and potential applications of engineering microstructures in thin films by FIBM rely on accurately detecting when sputtering of one layer is complete and milling of an underlying layer commences. Several methods for such end-point detection have been explored. For example, use of optical emission from sputtered neutral species and secondary ion mass spectrometry [3] are possible options for this task. Here, absorbed sample current is used as an indicator of end-point. For convenience of high sputter rates, and hence short milling times, 0.1 μm gold films deposited on silicon were selected for the end-point trials.

Rectangular apertures were milled through the gold film whilst monitoring the absorbed current. Figure 3 shows the typical variation of averaged absorbed sample current with time. Doubly charged gold ions were used. An array of 64×64 pixels was addressed and the absorbed sample current digitised at each pixel position. The beam dwell time at each pixel position was 1 ms. The average absorbed current (I) at the end of each scan was calculated and the ratio of this value to average initial current (Io) plotted. The value of Io was ≈ 0.5 nA. Figure 3 indicates that the gold film is removed after about 35 s. The rate of material removal is of order 1 $\mu m^3 s^{-1}$.

The control software was set up to mill five concentric circles into the gold film with radius of the outer circle 37.5 μm and radius difference of 7 μm. Each circular path was repeated 100 times. The result is shown in Figure 4(a). No attempt was made to correct for the increased ion dose due to closer pixel separation as the radius was reduced so consequently the depth of cut increases for the inner circles. Figure 4(b) shows the result of milling an annulus completely through the film, produced by scanning the ion beam in concentric circles and decrementing the radius by 0.15 μm. An outer radius of 30 μm and an inner radius of 20 μm was programmed.

By setting the inner radius to zero, through holes can readily be produced as shown in Figure 5(a) and Figure 5(b). The linear array of points above the hole mark the extremities of beam linescans during dimensional metrology trials, where dwell time was increased whilst transferring data. At the lower left corner sputtering at the position where the beam was 'parked' between machining operations is seen. Jitter due to noisy beam deflection voltages is evident. In Figure 5(b) a typical effect of inadequate mass separation is seen in the form of a ghost feature milled adjacent to the hole.

Conclusions

FIBM can be used for ultraprecision fabrication of microminiature three-dimensional structures. In-situ metrology utilising the same ion beam with suitable detectors is feasible. Apertures, blind holes and recessed holes with complex geometries can be fabricated. Such structures could find application as valve seatings, motor housings or channels in microelectromechanical systems.

Acknowledgments

The author thanks the Science University of Tokyo, Noda-shi, Chiba-ken 278, Japan, for provision of a Visiting Research Fellowship during the course of this work.

References

[1] J.Melngailis, J.Vac.Sci.Technol.B, **5**, 469-495 (1987).
[2] T.Suganuma, J.Electron Micros., **34**, 328-337 (1985).
[3] L.R.Harriott and M.J.Vasile, J.Vac.Sci.Technol.B, **7**, 181-187 (1989).

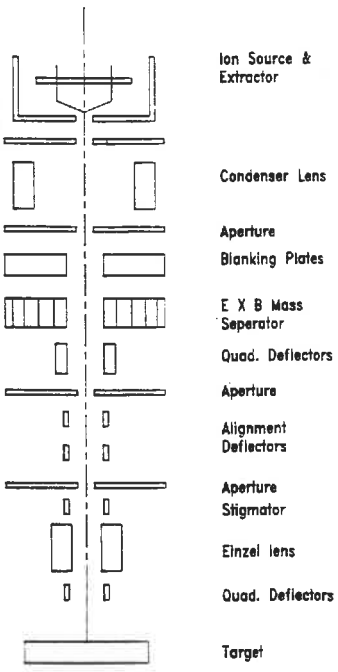


Figure 1. Schematic of FIB column.

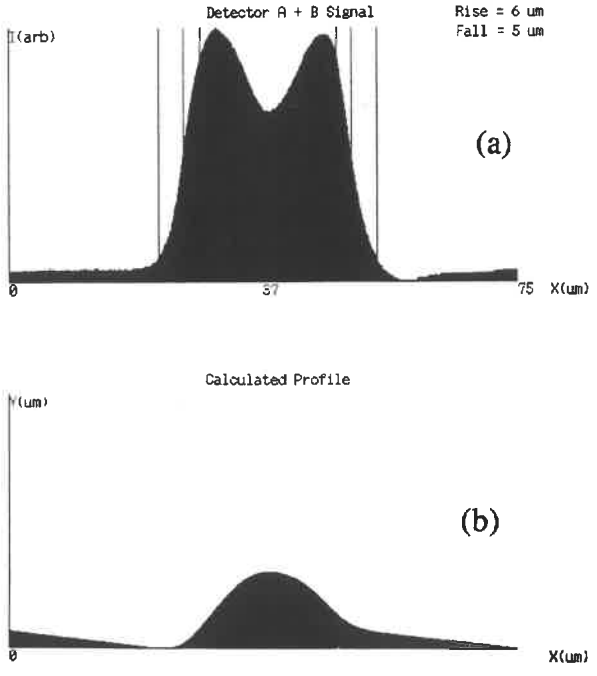


Figure 2. Linescan data (see text).

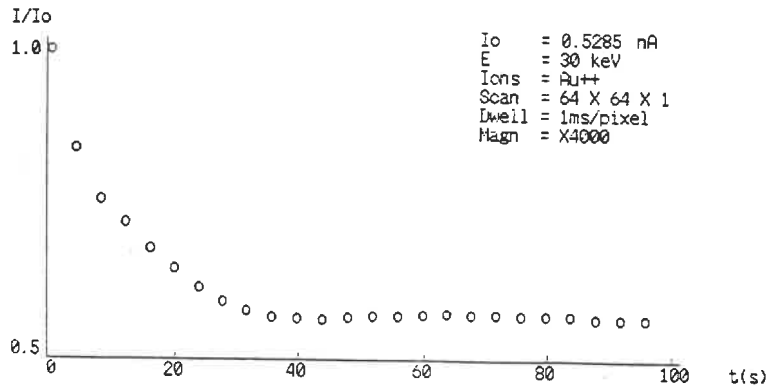


Figure 3. Averaged absorbed sample current as a function of time.

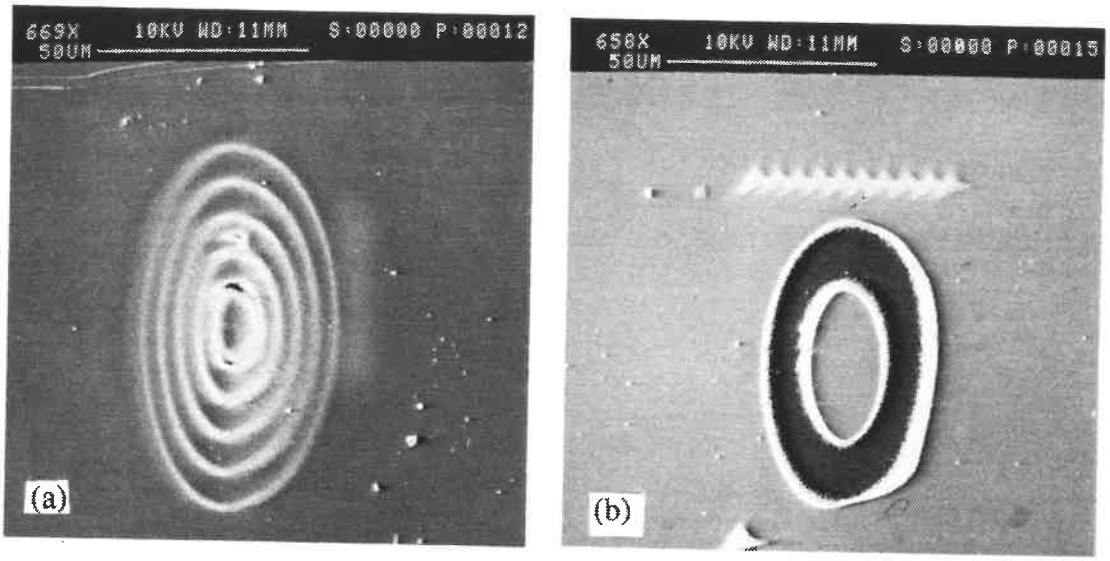


Figure 4. Concentric rings and annular aperture milled by FIB.

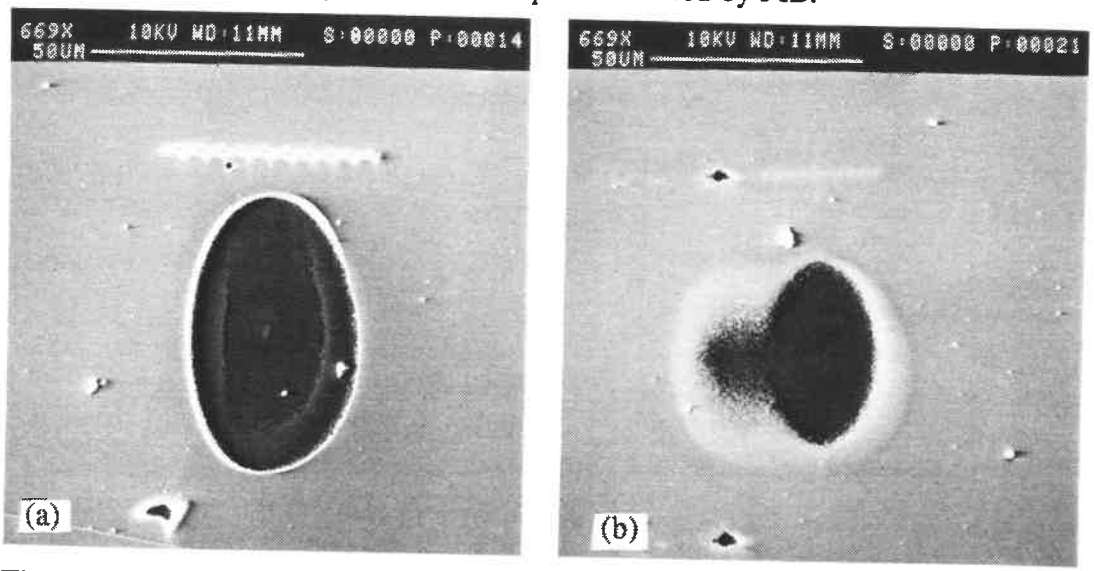


Figure 5. Apertures milled by FIB.

Environmental Effects on Subsurface Damage in Glass Machining

D. DePiero and T. G. Bifano
Aerospace and Mechanical Engineering
Ultraprecision Grinding Research Laboratory
Boston University
Boston, MA 02215

To establish a relationship between chemomechanical environment and material-removal processes in micro-grinding, a series of controlled scratching and fixed-abrasive grinding tests were performed on ULE glass (Corning Code 7971 Titanium Silicate Low Expansion Glass). ULE glass was also used in a recent study of chemomechanical effects in loose-abrasive microgrinding [1].

All scratching experiments were made using a Wilson Tukon microhardness tester with a Vicker's diamond pyramid indenter. The apparatus was modified to permit lateral scratching during the indentation loading cycle, using a motorized micrometer. After scratching at a pre-selected normal load and scratching speed, the scratched surfaces were examined using a video-microscope camera to archive the experimental conditions and nominal results on videotape. Scratches were subsequently analyzed using an SEM, then etched in 40% HF acid for 30 seconds and re-analyzed under SEM and an optical microscope. It was found that etching revealed cracks that were invisible in the unetched sample.

Environments tested included a homologous series of nine alcohols and de-ionized water. All scratches were performed at 5 gm normal load and 100 $\mu m/sec$ scratching speed. For each environment, 10-15 scratches were made. Figure 1 illustrates some typical scratch features before and after etching to reveal subsurface damage. Many scratches suffered subsurface damage, appearing in the following sequence of increasing damage: median cracks beneath the base of the scratch, radial cracks nearly perpendicular to the scratching direction, and lateral cracks extending nearly parallel to the sample surface from the sides or end of the scratch. To evaluate the fracture results, all scratches were categorized as either having subsurface damage (median, radial, or lateral cracks observed) or not having subsurface damage (purely ductile scratch). Figure 2 is a plot of the results. A correlation exists between subsurface damage and scratching environment. With the exception of hexanol, subsurface damage decreases steadily from the smaller molelcular weight alcohols through heptanol. The minimum in subsurface damage level corresponds to heptanol. This is the same environment that resulted in zero ζ-potential in Westwood's studies on glass machining [2]. Damage was most prevalent for ULE samples scratched in a water environment, corresponding to a large value of ζ-potential. Subsurface damage increases somewhat in octanol, and increases again in nonanol.

Further tests involved ultraprecision machining of ULE on a two-axis diamond grinding apparatus. This laboratory-scale grinder has a stiffness normal to the grinding contact of 50 MN/m and is controlled in real-time to a precision of 10 nm using piezoelectric actuation and capacitance gage sensing. It was found that the chemomechanical effect associated with $\zeta \approx 0$ (heptanol and octanol) was a transition from fracture-dominated brittle grinding to plastic-flow-dominated ductile grinding. SEM photomicrogaphs of surfaces ground in heptanol ($\zeta \approx 0$) nonanol ($\zeta < 0$) and water ($\zeta << 0$) illustrate the dramatic effect of environment on grinding damage (Figure 3). The specific grinding energy (inversely proportional to machining efficiency) also showed a significant dependence on environment. For the simple grinding geometry used in these experiments, in which the wheel

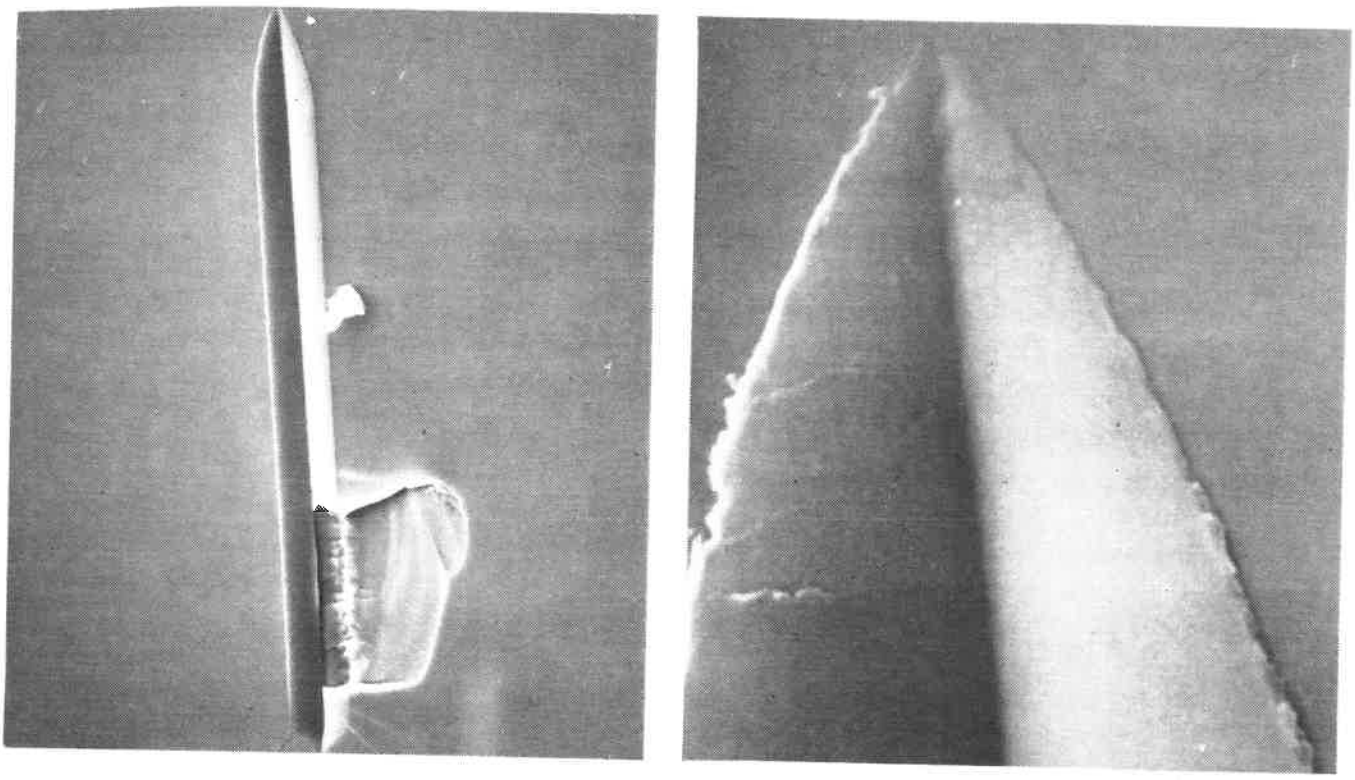


Figure 1: Typical features in single-point scratching of ULE. Left: unetched scratch with lateral and median cracks. Right: etched end of scratch with no subsurface damage.

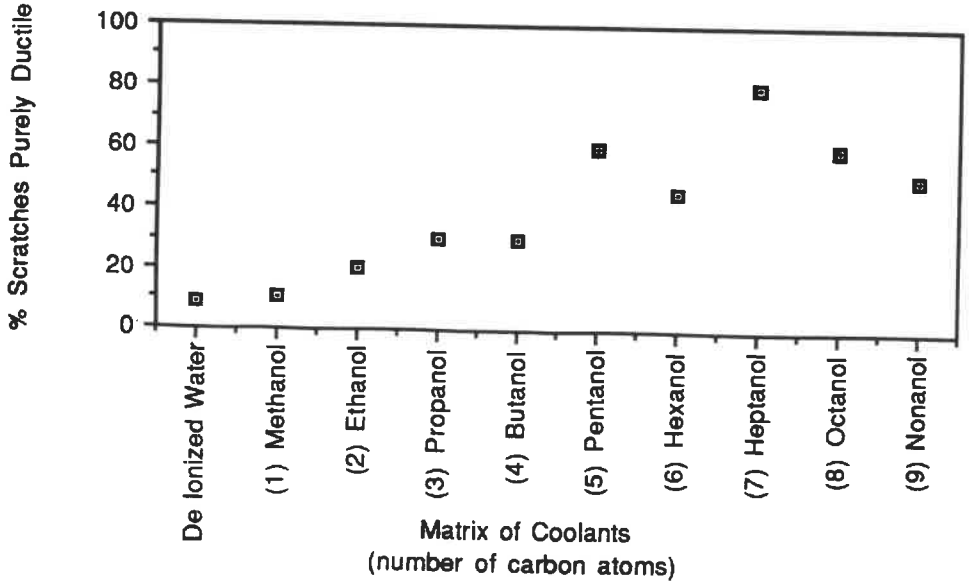


Figure 2: Subsurface damage in single-point scratching of ULE. Load: 5 gm. Speed: 100 $\mu m/s$

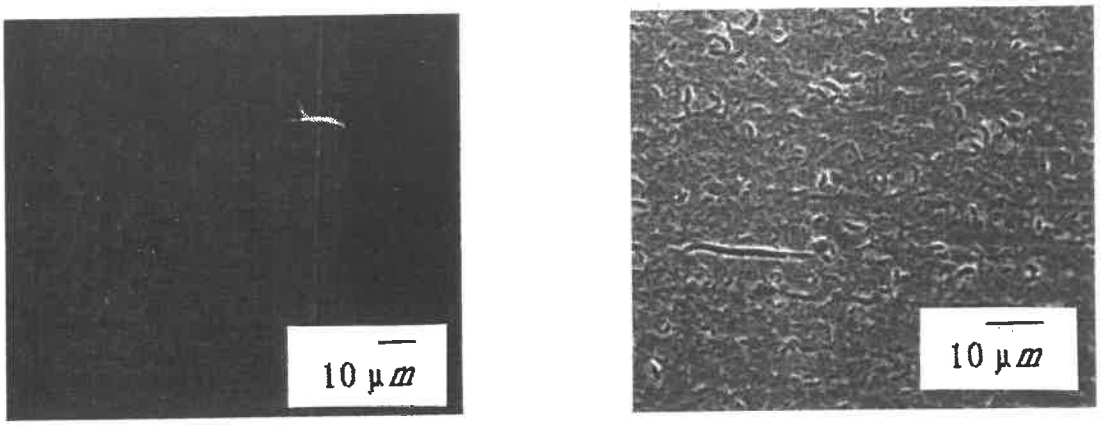


Figure 3: Subsurface damage in micro-ground ULE. Left: Heptanol environment. Right: Water environment.

speed is much greater than the crossfeed speed, the specific grinding energy is given by

$$\mu = \frac{F_t V_w}{MRR} \tag{1}$$

where F_t is the tangential grinding force, V_w is the grinding wheel peripheral speed, and MRR is the material removal rate. In water, nonanol, octanol, and heptanol environments, the specific grinding energy rose gradually with continued grinding, indicating wear of the abrasive grains with use. However, at a given overall grinding depth, μ was always considerably largest in an environment of heptanol followed by octanol, nonanonol, and finally water. At a grinding depth of 10 μm the specific grinding energy in water was $2.2 \times 10^{13}\ J/m^3$, in nonanol $2.8 \times 10^{13}\ J/m^3$, in octanol $5.6 \times 10^{13}\ J/m^3$, and in heptanol $9.5 \times 10^{13}\ J/m^3$.

These results (reduced machining efficiency and reduced subsurface damage in heptanol and octanol, with $\zeta \approx 0$) are consistent with Westwood's observations in cutting tests on glass with a non-cascading ball mill. Later work by Cuthrell on multipoint diamond machining of glass also showed a decrease in machining efficiency with increased surface hardness. A recent study by Golini and Jacobs using loose abrasives showed a brittle-ductile transition for ULE glass in heptanol and octanol environments. It is important to note that all of these experiments are performed near the ductile-brittle transition in machining rate. In very brittle-regime machining of glass by multipoint diamond tools, Westwood found that the correlation was reversed: an *increased* machining efficiency accompanied chemomechancial environments in which $\zeta \approx 0$. The fact that a particular environment can enhance efficiency and fracture at high machining rates (brittle-regime grinding) and then inhibit efficiency and fracture at low machining rates was demonstrated by Westwood in a series of experiments on polycrystalline alumina [2]. For diamond grinding of glass, the implications of this phenomenon are significant: the same environment can be used to enhanced material removal rate in coarse figuring operations and to inhibit fracture in finishing operations. This phenomenon was first observed in loose-abrasive grinding of glass by Phillips, Crimes, and Wilshaw [3].

Chemomechanical Toughening Hypothesis

Faced with empirical evidence of *increased* critical depths-of-cut (i.e. more fracture and lower specific grinding energy) associated with increased surface hardness ($\zeta \approx 0$) in microgrinding of glass,

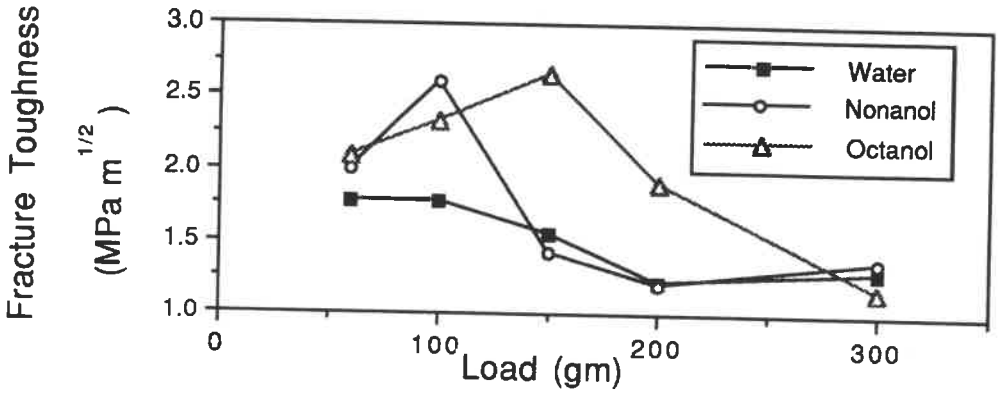


Figure 4: Fracture toughness as a function of load and environment, for ULE glass.

the ductile-regime grinding model ($d_c \propto E \cdot K_c^2/H^3$) suggests a hypothesis regarding chemomechanical effects: The fracture toughness K_c of the glass surface must also increase with a decrease in the absolute value of ζ-potential for glass micro-machining. From the ductile-regime grinding model, it is clear that chemomechanical surface toughening has the opposite effect of chemomechancial surface hardening on the critical grain-depth-of-cut.

Surface properties were measured in a series of microindentation tests to determine the validity of the chemomechanical toughening hypothesis. As predicted there is an unambiguous increase in the fracture toughness in octanol compared to water (Figure 4).

In ULE the critical-depth-of-penetration calculated using the microindentation data is twice as large in octanol than in water. This was qualitatively verified in constant-removal-rate micro-grinding experiments, where ULE glass ground in octanol was entirely damage-free, while ULE glass ground in water experienced 100% surface fracture. At larger loads, (simulating a grinding regime that would be entirely brittle) the value of K_c^2/H^3 is smaller in octanol than in water, suggesting that octanol would be a more eml rittling environment. This corresponds with the experimental results found bv Westwood for extreme brittle-regime multipoint grinding of glass.

Conclusions

Chemomechanical increases in the critical depth-of-cut can be used to facilitate a transition from brittle-regime grinding to ductile-regime grinding, without changing other machining conditions and without increased machine precision. This has been demonstrated through models and experiments in this paper.

References

[1] Golini, D., and Jacobs, S. D., **Appl. Optics,** [**30**]19, pp. 2761-2777, 1991.

[2] Westwood, A. R. C., and Mills, J. J., in *Surface Effects in Crystal Plasticity,* R. M. Latanision, J. T. Fourie, Eds., Noordhoff, Leiden, pp. 835-863, 1977.

[3] Phillips, K., Crimes, G. M., and Wilshaw, T. R.. **Wear,** [41], pp. 327-350, 1977.

Influence of Quality of Natural Diamond on Tool Wear in Precision Turning

Masashi Sawa*, Masami Masuda*, Yukio Maeda* and Ryuu Itoo**
* Production Engineering Research Laboratory, Hitachi, Ltd.
292 Yoshida-cho, Totsuka-ku, Yokohama 244, Japan
** Odawara Works, Hitachi, Ltd.
2880 Kouzu, Odawara 256, Japan

1. Introduction

When nonferrous metals are machined to mirror-like surfaces with single-crystal diamond tools, the cutting edges wear gradually by a submicron quantity with the increase of cutting distance. However, it has been empirically verified that the wear rate varies widely in a quantity of diamond tools, and this has become a hindrance in mass production machining process. It has been reported that the wear rate of tools differs according to the crystallographic orientation of diamonds[1], that it has some correlation with the nitrogen impurities in the diamonds[2),3)], and that there is some correlation between the microfracture strength and the infrared absorption coefficient of diamonds[4)]. Diamonds are essentially colorless and transparent, but many yellow and brown diamonds result from including of impurities. Tool manufacturers insist that hardness and toughness of diamond vary with the color, and the brown diamonds are harder and have a longer tool life than the yellow ones [5)].

This paper deals with effects that the various colors of diamonds exert on tool wear in order to elucidate the cause by which the tool life disperses.

2. Experimental Procedure and Conditions

A precision lathe was used for the experiments. The axial runout of the main spindle, composed of air bearings, is within 0.03 μm TIR. The yawing error of the hydrostatic feed table is within 0.02 μm/100 mm of tool feed travel. The cutting conditions are shown in Table 1, and an example of the tool shape and the profile measured with plunge-cut replica method are shown in Fig.1.

A Fourier transformation infrared spectrometer was used to measure the infrared absorption coefficient which are correlated to the nitrogen impurities in diamonds. After diamonds were ground into parallel-plate-shaped samples with 1 mm thick, the infrared absorption coefficient was measured. Subsequently, straight tools with a cutting edge of 1.2 mm long were fabricated, and the crystallographic orientation of these tools was detected by the Laue back-reflection process in which the faces of the tools were X-rayed vertically.

3. Experimental Results and Discussion

3.1 Comparison of the colors and the infrared absorption coefficients of diamonds

Natural diamonds are classified into four types Ia, Ib, IIa and IIb on the characteristics of electromagnetic wave absorption in the band from ultraviolet to infrared. Most of the natural diamonds belong to type Ia, and these diamonds are colored yellow, light yellow, white, light brown, brown and so on.

Though the correlation of such colors as yellow, light yellow, white, light brown and brown to the impurities in the diamonds is substantially unknown, a lot of evaluation has been tried regarding the nitrogen impurities in diamonds as determined by infrared absorption spectra.

Therefore, 20 diamonds with different colors were formed into parallel-plate-shaped samples, and their infrared absorption spectra were measured. A typical example of the spectra is shown in Fig.2. The absorption spectrum of wavelengths between 3 μm and 6 μm in the figure corresponds to the crystal structure of pure diamond, and the absorption in the wavelength range of 6 ~ 13 μm originated from the defects due to nitrogen impurities.

Ikawa reports that the microfracture strength of diamonds is correlated to the infrared absorption coefficient a1 at the wavelength of 7.3 μm among these defects due to nitrogen impurities[4)]. According to Wong [2)] and Syn[3)], the wear of the tools differs according to the ratio RA (= a1 / a2) of the infrared absorption coefficient a1 at the wavelength of 7.3 μm to that a2 at the wavelength of 7.8 μm in the absorption spectra due to nitrogen impurities.

These results are summarized in Fig.3. In this figure, the brown and light brown group had infrared absorption coefficients a1 and a2 as small as 1/10 ~ 1/2 in contrast with yellow and light yellow group, and the white diamond had intermediate values between those groups. When the calculation was made on the basis of Ikawa's report[4)], it was considered that the microfracture strength of brown diamonds (the mean value of a1 is 1.3 cm^{-1}) was about 1.2 times as high in comparison with yellow diamonds (the mean value of a1 is 17.4 cm^{-1}).

3.2 Selection of diamonds

In addition to the above microfracture strength, there have been also described the ratio RA of infrared absorption coefficients (= a1 / a2)[2),3)] and the crystallographic orientation[1)] of diamonds as factors which are considered to affect the wear of the tools. Here the authors tried to use tools which have almost the same crystallographic orientation and ratio of infrared absorption coefficients as one other.

First, the ratio RA of the infrared absorption coefficients of the diamonds in Fig.3 was determined as shown in Fig.4. In the figure, the following can be conveyed.

(1) The ratio RA of infrared absorption coefficients is in the range of 0.14 ~ 0.95.
(2) The RA values show the dispersion in the range of 0.3 ~ 0.6.
(3) In the yellow and brown diamonds, any significant differences are not observed.

Bearing in mind the above facts, typical one of 4 diamonds in each color group ware selected, as shown in Table 2. Their infrared absorption coefficient a1 and the ratio RA of infrared absorption coefficients are close to the mean values of each color group. These five diamonds representative of respective color groups implicate the following:

(1) The infrared absorption coefficient a1 of the yellow diamond is about 15 times as large as that of the brown one.
(2) The ratio RA of infrared absorption coefficients is in the range of 0.41 ~ 0.65, and the RA values of the yellow and brown diamonds are nearly the same.

These five diamonds were formed into straight tools with a straight cutting edge of 1.2 mm long (Fig.1). On the basis of the Kawakubo's report [1)], the crystallographic orientation on the flanks was set to the {310} plane. The angular errors from nominal crystallographic orientation were within 2 deg in the five tools.

3.3 Correlation of the colors of diamonds and the wear of tools

The wear of diamond tools showed a chamfered form composed of the {110} plane over the width of the tool feed per revolution. The authors have clarified that there is a strong correlation between the recession of the cutting edges δ1 shown in Fig.5 and the machined surface roughness. The change of the recession of cutting edges in the five tools with increase of cutting distance is shown in Fig.6. In the figure , the following can be described.

(1) The increasing rate in the recession of cutting edges (that is the wear rate of cutting tool) is nearly constant at 0.002 μm/km in the cutting distance exceeding 50 km in all the tools.

(2) Whereas the recession of cutting edges δ1 at the cutting distance of 160 km in brown cutting tool and light brown one is 0.5 μm, δ1 of yellow one, light yellow one and white one is as small as 0.4 μm.

The microphotographs of cutting edges and the transcripts of cutting edge profiles of the brown cutting tool and the yellow one at the cutting distance of 160 km in Fig.6 are shown in Fig.7. No abnormal wear such as boundary wear or chipping was observed on the worn parts of end cutting edges of both tools. From this fact, it can be described that the tools wore gradually in the normal cutting conditions.

With regard to the above fact, the wear rate of the tools was nearly constant in the steady wear region over the cutting distance of 50 km. However, since the wear rate of the cutting tool at the initial cutting stage is different, a difference of 0.1 μm at maximum arises in the recession of the cutting edges at the cutting distance exceeding 100 km. It is supposed that the microcracks which cannot be confirmed with a microscope have aroused on the cutting edges in the sharpening grinding process because brown diamonds are harder to grind than yellow ones, so the difference of initial wear rate is originated in accordance with the quantity of these microcraks.

3.4 Effect of tool wear on the machined surface roughness

The change of the machined surface roughness with the increase of cutting distance, which was obtained by the same experiment as that in Fig.6, is shown in Fig.8. According to the figure, the following can be described.

(1) The machined surface roughness in each tools is improved with the increase of cutting distance and becomes constant at 0.02 μm Rmax.
(2) In the brown tool and light brown one, of which the wear rate at the initial cutting stage is larger than others, the cutting distance until the machined surface roughness reached 0.02 μm Rmax is shorter than the other tools.
(3) The change of the machined surface roughness with the increase of cutting distance in the yellow tool followed nearly the same as brown one.

The authors have clarified previously the following[6)]:

(1) The machined surface roughness improved with the increase of the recession of cutting edges δ1 in the cutting at a positive tool setting angle.
(2) The reason of above fact is that a submicron stock was left on an irregular surface by the worn part of a tool, and the trailing cutting edge cut this uncut stock minutely; thus torn-off marks were removed.

As denoted by comparison with Fig.6, the machined surface roughness becomes 0.02 μm Rmax when the recession of the cutting edges δ1 reached 0.25 ~ 0.3 μm. Furthermore, the mechanism of improving the machined surface roughness due to the minute cutting accompanying the recession of cutting edges was verified.

4. Conclusions

As the results, within the range of these experimental conditions, it was clarified that :
(1) Some correlation was observed between the diamond colors and infrared absorption coefficients a1 and a2 at wavelengths of 7.3 μm and 7.8 μm, and in yellow diamonds, both a1 and a2 were about 10 times as large as those of brown ones,

(2) In the initial wear region at cutting distance of not more than 50 km, a difference of 0.1 μm at maximum arose in the recession of the cutting edges of the tools made from the diamonds colored yellow, light yellow, white, light brown and brown,
(3) In the steady wear region at cutting distance exceeding 50 km, any significant difference of the increasing rate of the recession on the cutting edges in relation to cutting distance was not observed between yellow and brown diamond,
(4) There was little difference in the change of the machined surface roughness in relation to cutting distance between yellow and brown diamond.

5. References

1) Y.Kawakubo, K.Morita: Tool Life of the Diamond Cutting Tool for Ultra-precision Machining, Proc. of Autumn Conf. of JSPE, (1982) 407 (in Japanese).
2) C.J.Wong: Fracture and Wear of Diamond Cutting Tools, Trans. ASME, 103, (1981) 341
3) C.K.Syn, J.S.Taylor and R.R.Donaldson: Diamond Tool Wear vs Cutting Distance on Electroless Nickel Mirrors, Proc. SPIE, 676, 8, (1986) 26.
4) N.Ikawa, S.Shimada and H.Tsuwa: Non-Destructive Strength Evaluation of Diamond for Ultra-Precision Cutting Tool, Ann. CIRP, 34, 1, (1985) 117.
5) H.Hamada: Ultra- Precision Cutting Tool, Jour. of JSPE, 36,6, (1984) 673 (in Japanese).
6) M.Masuda, Y.Maeda, T,Nishiguchi and M,Sawa: A Study on Diamond Turning of Al-Mg Alloy, Ann. CIRP,38,1, (1989) 111.

Table 1 Cutting conditions

Tool	Straight tool (Single crystal diamond)
Workpiece	Al-Mg alloy (annealed)
Cutting speed	7.5~16.5m/s
Feed rate	50μm/rev
Depth of cut	10μm
Cutting fluid	White kerosene
Tool setting angle	+0.016°

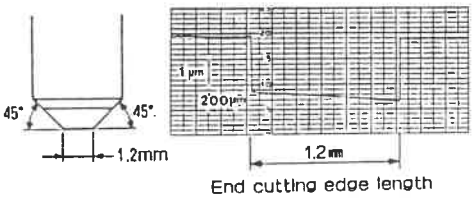


Fig. 1 Shape of a straight diamond tool

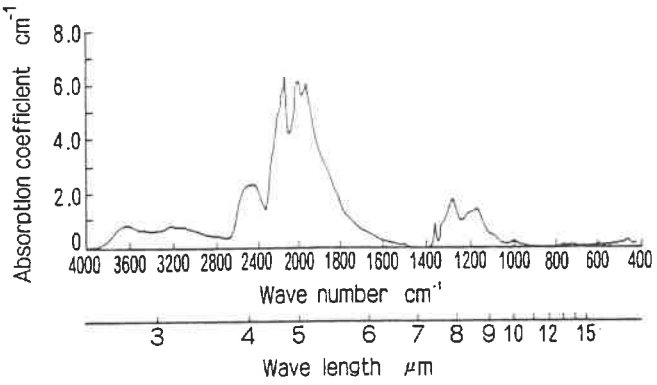


Fig. 2 Typical FT-IR spectra of natural diamond

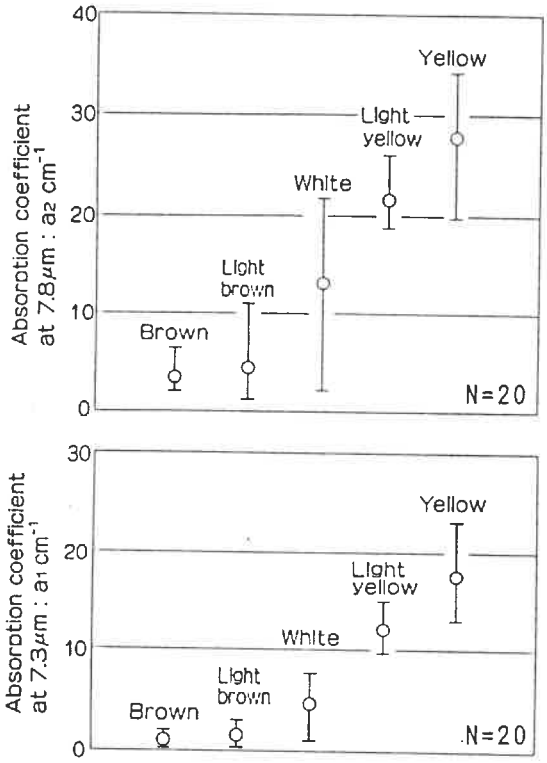


Fig. 3 Absorption coefficients at 7.3μm and 7.8μm in natural diamonds with various colours

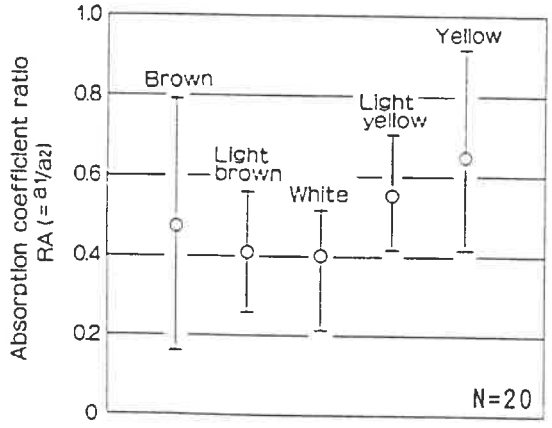
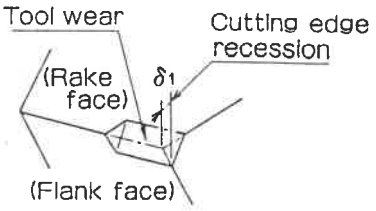


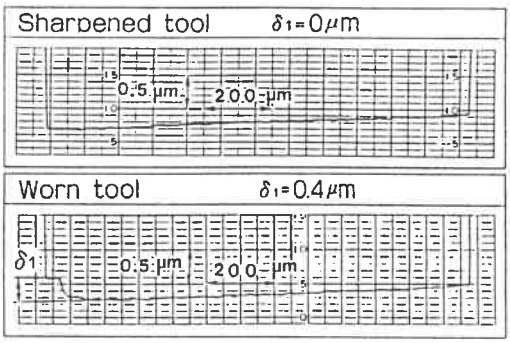
Fig. 4 Absorption coefficient ratio of natural diamonds with various colours

Table 2 IR absorption coefficient of natural diamonds used in experiment

Tool sample	Colour	Absorption coefficient (cm^{-1})		$RA=\frac{a_1}{a_2}$
		$a_1(\lambda=7.3\ \mu m)$	$a_2(\lambda=7.8\ \mu m)$	
A	Brown	1.3	2.0	0.65
B	Light brown	0.8	1.7	0.47
C	White	4.9	12.0	0.41
D	Light yellow	11.3	26.6	0.42
E	Yellow	19.8	34.4	0.58



(a) Wear model of straight diamond tool



(b) Cutting edge profiles of plunge-cut replica

Fig. 5 Cutting edge profiles of plung-cut replica with sharpened and worn tools

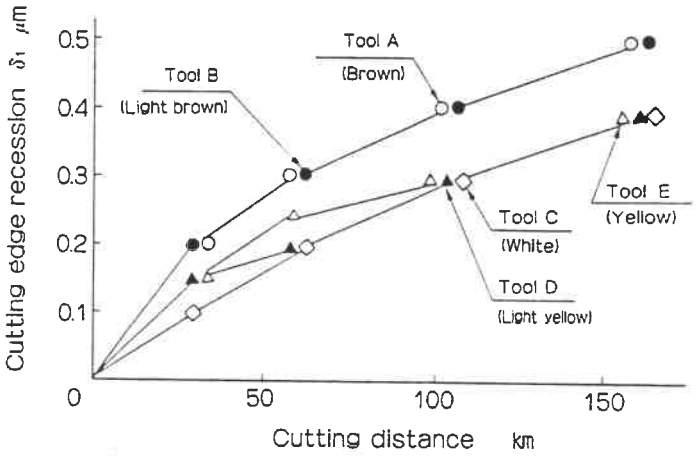


Fig. 6 Effect of cutting distance on cutting edge recession

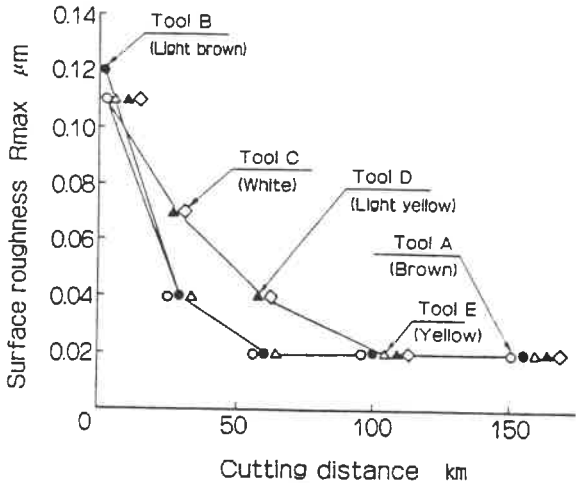


Fig. 8 Dependence of cutting distance on surface roughness

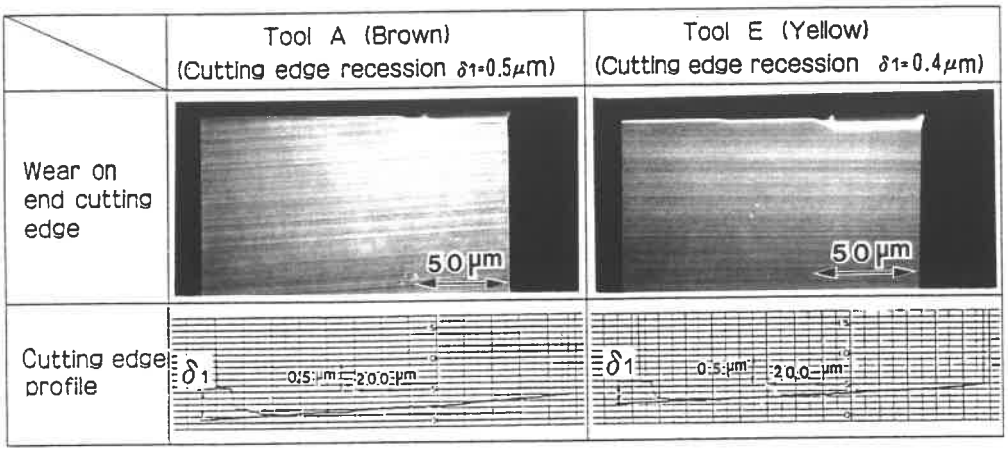


Fig. 7 Microphotographs and cutting edge profiles of worn tools

EVALUATION OF ENVIRONMENTALLY SAFE CLEANING AGENTS FOR DIAMOND TURNED OPTICS

Lonnie A. Theye, Robert D. Day
Los Alamos National Laboratory
Los Alamos, NM 87545

Rudolf Schubert
Bell Communications Research
Red Bank, NJ 07701

Background and Need:

Precision machining of metal surfaces using diamond turning has increased greatly in popularity at LANL in recent years. Similar techniques are used extensively to manufacture metal mirrors for use in laser applications. The diamond turned surfaces are easily damaged, making the selection of a cleaning agent very critical.

These surfaces have been traditionally cleaned using Trichloroethane (TCA) to remove residual oil remaining from the machining process. The TCA was then removed with an ethanol rinse, leaving a residue free surface. Recently, however, TCA was pronounced environmentally unsafe. Consequently, we are searching for an environmentally safe cleaning agent for these diamond turned metal optics.

The concern with using alternative solvents is the potential for residual surface films that produce reflectivity changes related to a combination of wavelength, surface coverage, film thickness and dielectric properties. Therefore, we have initiated a program for testing the effectiveness of a variety of environmentally safe solvents used to clean diamond turned optical surfaces.

Our basic test plan consists of comparing a number of environmentally safe solvents against the TCA/ethanol cleaning system. We have identified twelve candidate solvents, but have only been able to perform a partial test on one of them to date. This paper discusses the results obtained to date using this solvent known as P F [1].

Experimental Procedure:

Three different materials have been used for the tests. They are: Oxygen free high conductivity (OFHC) copper, electrolytic tough pitch (ETP) copper and 6061 aluminum. These materials were selected because they are currently used for many diamond machining applications. Test specimens were cut from a single piece of bar stock into blanks 3/4 inch in diameter and 1/8 inch thick. The surfaces to be tested were finished using a single-point diamond tool with a tool nose radius of approximately 3mm. The machining conditions were adjusted so that a 1 υm/rev feed was obtained.

A fixture was made to simultaneously machine nine blanks. Clean mineral oil was used as a cutting fluid. After removing the blanks from the machine they were stored in a mineral oil bath until they could be cleaned. The mineral oil plus any particulates generated by the machining process must be cleaned from the surface.

The cleaning procedure consisted of squirting cleaning solvent onto the oily surface and then air-drying with pressurized clean air. The fluid coming off of the surface during the cleaning process was collected for subsequent chemical analyses.

The following techniques are used to evaluate each solvent's effectiveness:

A. Bidirectional Reflectance Distribution Function (BRDF) scatter measurements
B. Ellipsometric analysis.
C. Observation of the surface using a Nomarski microscope.
D. Artificially aging for the equivalent of ten years using those solvents which provide the best surfaces based upon the three techniques listed above. After aging, an Auger analysis will be performed to determine surface contamination film thickness.

BRDF Measurements:
The BRDF is a measure of the amount of light scattered by a surface away from the normal angle of reflection. It is sensitive to both particulates and surface films and is a good measure of the surface performance in a typical application. For the purposes of this study, the BRDF measurements provide a functional criterion against which to evaluate the solvents. These measurements are made at a wavelength of 0.6328 μm and at off axis scattering angles between 0.5^{o} and 60^{o}. These correspond to a range of surface wavelengths between 0.73 μm and 73 μm.

Ellipsometry Measurements:
Ellipsometry is a sensitive technique used to measure surface film thicknesses of less than one monolayer. Therefore it is a very good quantitative method for determining how well a solvent removes surface oils. By using data obtained from the chemical analyses performed on the collected rinses along with a mathematical model, ellipsometric measurements can be used to determine the type and amount of surface contaminant. An advantage of this technique is that it works in air as opposed to vacuum.

Nomarski Microscope:
According to Bennett and Mattsson [2] the Nomarski (or Differential Interference Contrast) microscope is the most sensitive instrument for observing surface cleanliness next to the human eye. It is a good method for recording the surface condition on film, but since it produces a qualitative analysis it does not provide a numerical comparison of results.

Artificial Aging and Auger Analysis:
Long term performance of an optic is often an important requirement. Since the cleaning solvent affects surface chemistry, it was decided to artificially age the specimens after cleaning them. The aging is performed in a chamber with an atmosphere containing high concentrations of corrosive chemicals normally present in our environment. For example, a typical corrosive atmospheric mixture will be 10-ppb hydrogen sulfide, 10-ppb chloride, and 200-ppb nitrogen dioxide with a relative humidity of 70%[3]. The Auger analyses performed after aging determines the surface contamination thickness as a function of the solvents used.

Authors Index

Abler, Jeffrey A. 177
Aoyama, H. 194
Belak, J.F. 100
Bifano, Thomas G. 116, 200, 213
Blough, C. Gary 124
Bobroff, Norman 43, 157
Borchardt, B. 33
Bryanston-Cross, P.J. 169
Carr, J.W. 96
Chao, C-L. 112, 194
Chetwynd, D.G. 17
Chollet, Patrice 32
Clement, Andre 32
Clermont, Didier 32
Cumbo, M.J. 128
Cuttino, James F. 181
Davies, S.T. 209
Day, Robert D. 222
DePiero, D. 213
Dorion, T.D. 33
Dow, Thomas A. 62, 100, 104, 120, 181
Drescher, J.D. 100, 104
Duduch, J. 194
Elings, V.B. 153
Estler, W. Tyler 21, 54, 157
Evans, Chris J. 54, 93
Fawcett, Steven C. 120
Flynn, K. 194
Fujii, Teruo 149
Furutani, Akihiro 85
Futami, Shigeru 85
Garrard, Kenneth P. 62
Gee, A.E. 112, 194
Gerchman, Mark Craig 74
Geskin, E.S. 198
Gilsinn, David 21
Greier, Paul 157
Hara, Y. 194
Harary, Howard H. 32
Harb, S. 17
Hashimoto, H. 132
Hatamura, Yotaro 188
Henry, J. 33
Hercher, Michael 25
Hetrick, Paul 47
Hocken, Robert J. 50, 137, 149
Hosler, Joseph 200
Howard, Lowell P. 50
Hubbel, Peter I. 89
Ikawa, N. 100
Imai, K. 132
Itoo, Ryuu 217
Izraelev, Valentin M. 192
Jacobs, S.D. 128
Judge, T.R. 169
Kanai, Akira 80
Kanda, Torahiko 133
Kendall, Rodney A. 13
Kim, J. 96
Ko, K.L. 145
Komanduri, R. 100
Kramar, John 47
Kumaran, Senthill 149
Kunzmann, Horst 9
Lavigne, Claude 32, 137
Lin, T.Y. 145
Lindquist, A. 128
Loewen, Erwin G. 124
Lucca, D.A. 100
Luderich, J. 76
Maeda, Yukio 217
Mandak, J. 173
Mansfield, Daniel 141
Masuda, Masami 217
McCandlish, S.G. 194
Miller, Michele H. 62
Miller, Jimmie A. 137
Mitchell, Donald 25
Mitsuhashi, Masashige 133
Mitsuishi, Mamoru 188
Miyashita, Masakazu 80
Mizuno, H. 194
Moore, Duncan T. 124
Moriwaki, T. 100
Nagao, Takaaki 188
Narayan, C. 96
Ng, S.B. 128
Ohura, Toshio 165
Okuda, K. 100
Oldham, Nile 47
Penzes, W. 47
Phillips, S.D. 33
Pieczulewski, Dan 21
Polvani, Robert S. 93
Puttick, K.E. 194
Queen, Yun Hsia 21, 157
Rhorer, R.L. 100
Rich, T. 128
Ro, Paul I. 89, 177
Rodgers, Mark R. 153
Rudman, M. 194
Sabharwal, Y. 128
Sakaya, Katsuaki 165
Sawa, Masashi 217
Scattergood, Ronald O. 108, 205
Schaefer, John P. 70
Schellekens, P.H. 37
Schubert, Rudolf 222
Scott, Paul J. 161
Shimada, S. 100

Shimanovich, M.A. 198
Slocum, Alexander H. 66, 185
Smith, S.T. 17
Smith, Stanley M. 205
Soons, J.A. 37
Stedman, Margaret 29
Stevens, David M.G. 58
Stout, K.J. 145
Stowers, I.F. 100
Sullivan, P.J. 145
Takeda, J. 132
Tanaka, H. 100
Taylor, Lauren W. 62
Teague, Clayton 47
Theye, Lonnie A. 222
Tidwell, M. 108
Trumper, David L. 66
Weck, M. 76
Wheatley, T. 47
Wyntjes, Geert 25
Yi, Yuan 116
Yoshioka, Junichi 80